

Critical Review

New Approach Methodologies for the Endocrine Activity Toolbox: Environmental Assessment for Fish and Amphibians

Constance A. Mitchell,^{a,*} Natalie Burden,^b Mark Bonnell,^c Markus Hecker,^d Thomas H. Hutchinson,^e Magdalena Jagla,^c Carlie A. LaLone,^f Laurent Lagadic,^g Scott G. Lynn,^h Bryon Shore,^c You Song,ⁱ Sara M. Vliet,^j James R. Wheeler,^k and Michelle R. Embry^a

^aThe Health and Environmental Sciences Institute, Washington, DC, USA

^bNational Centre for the 3Rs (NC3Rs), London, United Kingdom

^cEnvironment and Climate Change Canada, Ottawa, Canada

^dToxicology Centre and School of the Environment & Sustainability, University of Saskatchewan, Saskatoon, Canada

^eReckitt R&D, Hull, United Kingdom

^fOffice of Research and Development, Great Lakes Toxicology & Ecology Division, US Environmental Protection Agency, Duluth, Minnesota

^gResearch and Development, Crop Science, Environmental Safety, Bayer, Monheim am Rhein, Germany

^hOffice of Pesticide Programs, US Environmental Protection Agency, Washington, DC

ⁱNorwegian Institute for Water Research, Oslo, Norway

^jOffice of Research and Development, Scientific Computing and Data Curation Division, US Environmental Protection Agency, Duluth, Minnesota

^kCorteva Agriscience, Bergen op Zoom, The Netherlands

Abstract: Multiple in vivo test guidelines focusing on the estrogen, androgen, thyroid, and steroidogenesis pathways have been developed and validated for mammals, amphibians, or fish. However, these tests are resource-intensive and often use a large number of laboratory animals. Developing alternatives for in vivo tests is consistent with the replacement, reduction, and refinement principles for animal welfare considerations, which are supported by increasing mandates to move toward an “animal-free” testing paradigm worldwide. New approach methodologies (NAMs) hold great promise to identify molecular, cellular, and tissue changes that can be used to predict effects reliably and more efficiently at the individual level (and potentially on populations) while reducing the number of animals used in (eco)toxicological testing for endocrine disruption. In a collaborative effort, experts from government, academia, and industry met in 2020 to discuss the current challenges of testing for endocrine activity assessment for fish and amphibians. Continuing this cross-sector initiative, our review focuses on the current state of the science regarding the use of NAMs to identify chemical-induced endocrine effects. The present study highlights the challenges of using NAMs for safety assessment and what work is needed to reduce their uncertainties and increase their acceptance in regulatory processes. We have reviewed the current NAMs available for endocrine activity assessment including in silico, in vitro, and eleutheroembryo models. New approach methodologies can be integrated as part of a weight-of-evidence approach for hazard or risk assessment using the adverse outcome pathway framework. The development and utilization of NAMs not only allows for replacement, reduction, and refinement of animal testing but can also provide robust and fit-for-purpose methods to identify chemicals acting via endocrine mechanisms. *Environ Toxicol Chem* 2023;42:757–777. © 2023 The Authors. *Environmental Toxicology and Chemistry* published by Wiley Periodicals LLC on behalf of SETAC.

Keywords: Adverse outcome pathway; Endocrine-disrupting compounds; In vitro toxicology

This article includes online-only Supporting Information.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

* Address correspondence to cmitchell@hesiglobal.org

Published online 15 February 2023 in Wiley Online Library (wileyonlinelibrary.com).

DOI: 10.1002/etc.5584

INTRODUCTION

Endocrine pathways regulate growth, development, reproduction, metabolism, and tissue function in vertebrates (US Environmental Protection Agency [USEPA], 2022b). Examples of endocrine pathways important for the functions listed are estrogen, androgen, thyroid, and steroidogenesis modalities,

which are the focus of regulatory programs such as the USEPA's Office of Pesticide Programs (OPP). A chemical that alters the function of an endocrine system and causes subsequent adverse effects in an intact organism, its progeny, or (sub)populations is referred to as an endocrine-disrupting chemical (EDC; International Programme on Chemical Safety, 2002). Endocrine activity, by contrast, describes an interaction with the endocrine system observed during a toxicological study that may or may not lead to an adverse effect. There is also scientific consensus that the World Health Organization (WHO)/International Programme on Chemical Safety definition does not necessarily mean that the adverse effect has to be demonstrated in an intact test animal but may be shown in adequately validated alternative test systems predictive of adverse effects in humans and/or wildlife (Solecki et al., 2017). Endocrine-disrupting chemicals may alter the endocrine system by mimicking endogenous hormones (e.g., binding to receptors) or altering the synthesis, metabolism, and transport of hormones (Schneider et al., 2019) and are of concern for both human and ecological health. Testing requirements related to identification of EDCs are developing worldwide but vary in approaches. These range from test requirements for all substances being registered or marketed for a particular use to exposure- and production volume-based scenarios to specific testing triggered by emerging concerns (Table 1).

Initial testing for endocrine activity typically involves short-term *in vivo* and *in vitro* assays used to evaluate the potential for interactions of a chemical with select endocrine pathways. Where activity is identified to be of concern, higher-tier *in vivo* testing is subsequently required to confirm the activity and establish adversity. Multiple *in vivo* test guidelines have been validated for mammals, amphibians, and fish by the Organisation for Economic Co-operation and Development (OECD) and the USEPA (OPPTS 890 test guidelines; Burden et al., 2022; OECD, 2021a). These guidelines often require the use of a large number of laboratory animals (Russell & Burch, 1959). Because of these animal welfare and economic concerns, there are mandates to move toward an "animal-free" testing paradigm worldwide (Frank R. Lautenberg Chemical Safety for the 21st Century Act, 2016; Oziel, 2017; Wheeler, 2019).

The potential for expanding the use of new approach methodologies (NAMs) in endocrine activity testing and assessment

New approach methodologies refer to "any technology, methodology, approach, or combination thereof that can be used to provide information on chemical hazard and risk assessment that avoids the use of [intact] animals" (OECD, 2005). However, the term NAM can also include whole-organism assays using some invertebrates or the nonprotected elutheroembryonic life stages of vertebrates (e.g., fish and amphibians, typically prior to independent feeding; Embry et al., 2010; von Hellfeld et al., 2020). For the purposes of the present study, we refer to NAMs as all (eco)toxicological tools

other than traditional *in vivo* toxicity tests. The USEPA (2021) similarly defines NAMs as "any technology, methodology, approach, or combination that can provide information on chemical hazard and risk assessment to avoid the use of animal testing." Generally, NAMs can provide information on molecular changes, molecular/cellular perturbations, toxicokinetics, toxicodynamics, and toxicity mechanisms rather than measuring apical endpoints that are assessed *in vivo*. New approach methodologies hold great promise to increase the efficiency of screening testing to inform regulatory decisions, while having the added benefit of reducing the need for more resource-intensive studies which are reliant on large numbers of animals (OECD, 2005). Currently, however, there is limited confidence in using these tools broadly for hazard and risk assessment (Burden et al., 2022); and to date NAMs have been primarily used for chemical screening, classification, or prioritization or to assist in providing a mechanistic understanding of effects observed in *in vivo* studies. This is partly because many NAMs lack the complex biological processes that *in vivo* systems possess, such as absorption, distribution, metabolism, and excretion (ADME) processes or cross talk among multiple tissues or organs. Furthermore, insufficient mechanistic knowledge of the biochemical pathways of interest may pose issues concerning the use and interpretation of NAM data. This collaborative effort strives to summarize the current state of the science regarding the use of NAMs to identify and predict chemical-induced endocrine effects and aims to identify steps needed to increase acceptance of NAMs in regulatory processes.

Adverse outcome pathways as a framework to support NAM development and application

The adverse outcome pathway (AOP) framework has been developed to support a more mechanistically focused approach to toxicity testing and drive the development and integration of toxicodynamic NAMs in safety assessment. Adverse outcome pathways are constructed using available information to define the sequence of molecular and cellular events required to produce an adverse effect on individuals, with potential consequences at the population level when organisms are exposed to a substance. Adverse outcome pathways entail linking a molecular initiating event, which describes the initial interaction of a chemical with a biomolecule, through a cascade of downstream key events (KEs) occurring at multiple levels of biological organization (i.e., cellular, tissue, organ), to an adverse outcome of regulatory concern (Ankley et al., 2010). The strength of the causal relationships between KEs (termed KE relationships) can be evaluated by weight-of-evidence (WoE) considerations (Collier et al., 2016), integrating evidence streams along the AOP. Changes to an organism induced by perturbations of the pathway can then be mapped along the AOP. With sufficient understanding and development of quantitative endocrine AOPs, NAMs can be developed and used in place of animal toxicity tests to determine whether a substance induces early KEs within the AOP and ultimately help to predict if an adverse outcome is likely. The estrogen,

TABLE 1: Current regulatory requirements and opportunities for new approach methodologies related to endocrine-disrupting chemicals

Region	Substance type	Legal basis	Requirements for endocrine testing using fish or amphibian models	Opportunities for NAMs
Europe	Plant protection products—active substances	Commission regulation (EU) 2018/605 amending Annex II to Regulation (EC) No. 1107/2009 by setting out scientific criteria for the determination of ED properties	OECD Conceptual Framework (OECD, 2018a) toolbox adopted. According to the Commission Delegated Regulation (EU) No. 2017/21003 and Commission Regulation (EU) No. 2018/605, the conclusions as to whether the ED criteria are met need to be drawn separately with respect to humans and nontarget organisms. The assessment strategy aims at making the most efficient use of the available data set to reach a conclusion. It is recommended to strive for a conclusion on the ED properties regarding humans and, in parallel, using the same database, to strive for a conclusion on mammals as nontarget organisms. Only where, based on this assessment, the criteria are not met for mammals as nontarget organisms would the assessment need to proceed by considering fish and amphibians because these are the taxa where standardized test methods and knowledge on how to interpret the results are available. Information on other taxa (e.g., birds and reptiles) should be considered if available (ECHA & EFSA, 2018).	Assessment of the mammalian data set includes the following in vitro assays: estrogen: OECD TG 455 (OECD, 2021c), OPPTS 890.1300 (USEPA, 2009d), and ToxCast ER Bioactivity Model Data; androgen: OECD TG 458 (OECD, 2020); steroidogenesis: OECD TG 456 (OECD, 2011a), OPPTS 890.1550 (USEPA, 2009e); EU B.57, OPPTS 890.1200 (USEPA, 2009c). In ECHA & EFSA (2018), eleutheroembryo assays (XETA, EASZY, RADAR assay) mentioned as Level-3 assays relevant for the evaluation of endocrine activity.
	Biocidal products—active substances	Commission Delegated Regulation (EU) 2017/2100 setting out scientific criteria for the determination of ED properties pursuant to Regulation (EU) No. 528/2012	Specific studies in fish may include a MEOGRT (OECD, 2015a) or a fish life-cycle study covering all the “estrogen, androgen or steroidogenesis-mediated” parameters foreseen to be measured in OECD TG 240. Specific studies in amphibians may include a LAGDA (OECD, 2015b). These studies do not need to be performed if endocrine activity is sufficiently investigated (i.e., a test according to OECD TG 229 [OECD, 2012a], 230 [OECD, 2009a], and 231 [OECD, 2009b] is available) and there is no indication that the substance has endocrine activity or effects potentially related to endocrine activity.	
	Industrial chemicals	Regulation (EC) No. 1907/2006 concerning REACH	EDs identified as substances of very high concern or equivalent level of concern are included in the candidate list and are subject to authorization/restriction processes. Currently (2022), proposals for standard information requirements are being discussed.	

(Continued)

TABLE 1: (Continued)

Region	Substance type	Legal basis	Requirements for endocrine testing using fish or amphibian models	Opportunities for NAMs
	Human pharmaceuticals	Regulation (EC) No. 726/2004	Currently (2022), the European Medicines Agency's Committee for Medicinal Products for Human Use guideline (EMA/CHMP, 2006) specifies that environmental risk of certain compounds, including endocrine-active ones, needs to be addressed irrespective of exposure. Some outlined guidance on testing approaches is also available in the accompanying Q&A document. A draft revision of the guidance released in 2018 retains the requirements for tailored risk assessment for endocrine-active substances and makes more specific recommendations for mechanism of action—specific testing, including OECD TG 229 (OECD, 2012a), 230 (OECD, 2009a), 234 (OECD, 2011b), 240 (OECD, 2015a), and 241 (OECD, 2015b).	
	Veterinary pharmaceuticals	Regulation (EC) No. 2001/82/EC	Required studies do not currently include studies specifically designed to identify endocrine-mediated effects.	Information from mechanism of action, mammalian toxicology, or open literature studies can inform requests for additional studies.
United States	Pesticides, pesticidal formulation inert chemicals, and environmental contaminants found in drinking water to which a substantial population is exposed	Food Quality Protection Act 1996; EDSP 1998 <i>Federal Register Notices</i> <ul style="list-style-type: none"> • August 11, 1998: EDSP • December 28, 1998: EDSP Statement of Policy 	EDSP uses a two-tiered approach: <ul style="list-style-type: none"> • Tier I <ul style="list-style-type: none"> • OPPTS 890.1100: Amphibian metamorphosis assay (USEPA, 2009f) • OPPTS 890.1350—Fish short-term reproduction assay (USEPA, 2009g) • Tier II <ul style="list-style-type: none"> • OPPTS 890.2200: MEOGRT (USEPA, 2015b) • OPPTS 890.2300: LAGDA (USEPA, 2015c) The 52 pesticides on List 1 have undergone Tier I screening, and no test orders for Tier II testing have yet been issued. List 1 includes chemicals that the USEPA selected based on exposure potential.	The USEPA is currently focusing on the pivot strategy of developing and implementing NAMs as alternatives. The goal is to develop a set of “nonanimal” high-throughput assays and computational bioactivity models as alternatives for all the assays in the current Tier 1 screening battery. In vitro consensus approaches have been proposed by the USEPA for ER (USEPA, 2015a) and AR activity (USEPA, 2017; “ER/AR pathway models”).
	Human pharmaceuticals	US: National Environmental Policy Act of 1969 Code of Federal Regulations Part 25—Environmental Impact Considerations	The US FDA has issued both a general guidance for environmental risk assessment (FDA, 1998) and a Q&A for endocrine-active substances (FDA, 2016). Neither of these request specific studies, but applicants are encouraged to request to consult the agency prior to submission.	

(Continued)

TABLE 1: (Continued)

Region	Substance type	Legal basis	Requirements for endocrine testing using fish or amphibian models	Opportunities for NAMs
Japan	Chemical substances	Japan Chemical Substance Control Law (2011) Current strategic program EXTEND 2017	<p>The Japanese strategy employs a two-tiered approach*:</p> <p>Tier I</p> <ul style="list-style-type: none"> • Medaka ERα reporter gene assay • Medaka ARβ reporter gene assay • OECD TG 231 (OECD, 2009b) • Fish screen (i.e., OECD TG 229 [OECD, 2012a] or 230 [OECD, 2009a]) • Juvenile medaka anti-androgen screening assay (under development) <p>Tier II</p> <ul style="list-style-type: none"> • OECD TG 241 (OECD, 2015b) • OECD TG 240 (OECD, 2015a) 	<ul style="list-style-type: none"> • Medaka ER a and b reporter gene assays • <i>Xenopus tropicalis</i> thyroid receptor b reporter gene assay
Canada	Pesticides		<p>The Pest Management Regulatory Agency considers toxicological studies that investigate the ED potential of pesticides as part of pesticide evaluations. For chemicals that exhibit ED properties or that induce adverse biochemical and/or cellular changes, the agency considers effect(s) resulting in a measurable holistic endpoint (e.g., mortality, reproduction, growth development, behavioral) to assess risk to nontarget organisms at environmentally relevant concentrations. It does not currently require studies evaluating specific ED effects for the premarket registration of new chemicals under the Pest Control Products Act, but available endocrine data are considered in risk assessments.</p>	<p>The Pest Management Regulatory Agency is open to considering any applicant rationale incorporating NAM results as a hazard flag for pesticides. It does not routinely use or receive NAM data for regulatory purposes. If NAM results were to flag an ED concern and exposure is expected, the substance could be considered as a candidate for additional data gathering. New approach methodology studies can impart valuable insight into the mechanistic toxicity of a chemical at the molecular and cellular levels; however, these effects may not necessarily translate into observable adverse effects at the organism level. Toxicity results of a mechanistic nature, therefore, are typically used qualitatively in the agency's risk assessments.</p>
	Chemical substances		<p>Endocrine tests are not presently a regulatory requirement under the Canadian Environmental Protection Act, 1999; but available endocrine data are considered and can be a driving factor in risk assessments. Safety, quality, and efficacy of pharmaceuticals are evaluated under the Food and Drugs Act.</p>	<p>Modes and mechanisms of action considerations, including endocrine-active substances, have been applied to ecological assessments (Okonski et al., 2021). The specific role for a given NAM will dictate both the types of data and the level of confidence needed to support its integration. For example, as replacements for apical, in vivo endpoints, NAMs would need to generate reliable, quantitative values to allow the estimation</p>

(Continued)

TABLE 1: (Continued)

Region	Substance type	Legal basis	Requirements for endocrine testing using fish or amphibian models	Opportunities for NAMs
				of predicted-no-effect concentrations, while NAMs intended to flag specific modes of action can be integrated through the application of additional assessment factors during predicted-no-effect concentration derivation. Canada's Ecological Risk Classification of organic substances approach under the Chemicals Management Plan utilizes WoE approaches which incorporate in silico flags.

NAM = new approach methodology; EC = European Commission; ED = endocrine disruptor/disruptor; OECD = Organisation for Economic Co-operation and Development; ECHA = European Chemicals Agency; EFSA = European Food Safety Authority; XETA = Xenopus elutheroembryonic thyroid assay; EASZY = endocrine active substances, acting through estrogen receptors, using transgenic tg(cyp19a1b::GFP) zebrafish embryos; RADAR = rapid androgen disruption adverse-outcome reporter; MEOGRT = medaka extended one-generation reproduction test; TG = test guideline; LAGDA = larval amphibian growth and development assay; REACH = Registration, Evaluation, Authorisation and Restriction of Chemicals; Q&A = question and answer; EDSF = Endocrine Disruptor Screening Program; OPPTS = Office of Prevention, Pesticides and Toxic Substances; USEPA = US Environmental Protection Agency; ER = estrogen receptor; AR = androgen receptor; FDA = US Food and Drug Administration; EXTEND = Extended Tasks on Endocrine Disruption; WoE, weight of evidence.

Source: Modified from Burden et al. (2022).

androgen, thyroid, and steroidogenesis modalities are of interest in the current regulatory paradigms. These pathways tend to be conserved across species, are relatively well studied, and in some cases have a robust set of in vivo adverse outcome data. Figure 1 maps the different methods and approach types that are available to support general chemical assessment for ecological endpoints along an AOP. Such organizational thinking can be applied for chemicals acting on endocrine pathways. However, the extrapolation from individual-based adverse outcomes to a population-relevant apical effect, in line with the protection goals of most environmental assessments, remains a significant challenge (Crane et al., 2019; Lagadic et al., 2020; Perkins et al., 2019).

It is worth noting that while NAMs hold promise for high-throughput screening and assessment with the added benefit of reducing dependence on traditional animal tests, they are not necessarily a one-to-one replacement. However, their data should be considered a key component in an integrated WoE approach for decision-making. A WoE approach uses a combination of information from several sources, with weight given to the available evidence based on expert judgment or reliability factors (e.g., quality of the data, consistency of results, and relevance of the information; European Chemicals Agency, 2016). Despite their limitations, NAMs provide the possibility to incorporate mechanistic information into the chemical assessment process, to increase the efficiency of such assessments, and to ultimately reduce the overall numbers of animals used in ecotoxicology testing and assessment. In particular, in silico, in vitro, and elutheroembryo models are more rapid and higher-throughput than traditional in vivo approaches, allowing more chemicals to be screened with fewer resources and in less time.

Aims of our review

In a collaborative effort led by the Health and Environmental Sciences Institute (HESI) and the United Kingdom's National Centre for the 3Rs, experts from government, academia, and industry met in early 2020 to discuss the current challenges of endocrine testing for fish and amphibians (Burden et al., 2022). In continuing the cross-sector initiative, the present study outlines the current state of the science for evaluating the endocrine-disrupting potential of chemicals in amphibians and fish using NAMs across the estrogen, androgen, thyroid, and steroidogenesis pathways. In the present study, we discuss the use of NAMs, their current and potential roles for ecotoxicology testing, and their potential to be used in paradigms like AOPs and integrated approaches to testing and assessments (IATAs). We also discuss the challenges of using NAMs for endocrine assessments and what is needed to reduce the uncertainties with such approaches. Movement toward the development, validation, and utilization of NAMs will lead to the replacement, reduction, and refinement of animal use needed for endocrine activity assessments while increasing the robustness and fitness for purpose of regulatory EDC assessments.

USING PROBLEM FORMULATION TO ENSURE THAT NAMs ARE FIT FOR PURPOSE

Problem formulation is defined as a systematic approach that identifies the factors critical to a specific assessment and considers the purpose, scope, and depth of the necessary analysis, analytical approach, available resources and outcomes, and overall management goal (Solomon et al., 2016). Because management goals may vary by sector, legislation, and geographic region, the specific regulatory context must be taken into account. One example of this is the difference between approaches to the regulation of EDCs in Canada, the United States, and Japan, where a risk-based approach is used, and in the European Union, where a hazard-based approach is employed (i.e., a lack of consideration of exposure in the assessment; Burden et al., 2022). Therefore, data and assessment methodology will vary significantly depending on the stated purpose, making it crucial that NAMs are evaluated to ensure that they are fit for purpose for the assessment type (e.g., prioritization, classification, hazard assessment, risk assessment). Regional regulations and current acceptance of NAMs under those regulations are listed in Table 1. The OECD Guidance Document 150 provides information on standardized test guidelines for evaluating chemicals for endocrine disruption (ED) and describes what tests are available to evaluate endocrine activity and ED (OECD, 2018a). While it has been used as a basis for some regulatory guidance, it does not in itself define a testing strategy. It should be recognized that although regulatory testing requirements may vary globally, many chemicals are intended for use worldwide, and the relevance of any test across geographies (NAM or otherwise) should be a key consideration. Streamlined data packages and harmonized data templates that satisfy multiple global requirements will ensure that collected data can be broadly utilized for regulatory purposes, thereby reducing duplicative testing and resource costs,

while still enabling marketing internationally (Sewell et al., 2017) and potentially reducing the need for animal testing.

BUILDING CONFIDENCE IN NAM DEVELOPMENT, EVALUATION, AND APPLICATION

Although NAMs can be used to reduce the number of animals and resources used in toxicity testing (Parish et al., 2020), *in vivo* data have been typically thought to provide more realistic information on chemical hazard, especially with respect to complex biological processes and systems within the assays (like ADME). Therefore, there are still several issues to overcome before there is sufficient confidence in NAM data for these tools to be incorporated more widely in safety assessment. There is a concern for high false-negative or false-positive outcomes with NAMs, leading to a reluctance to make decisions solely based on results from these assays. However, *in vivo* tests may also produce false negatives because their sensitivities and specificities for detecting endocrine activity have been shown to be not as high as those of some NAMs. A separate project from this collaboration will critically review the robustness of currently available *in vivo* tests. Building confidence in negative outcomes from NAMs is essential to avoid unnecessary animal studies. Confidence can be increased by using multiple NAMs covering multiple pathways to ensure that there are no “gaps” in coverage. Some degree of redundancy (i.e., several NAMs covering the same molecular initiating event) may also help in consolidating a positive or negative outcome. Next, to gain regulatory acceptance, NAMs must be “validated” (appropriate definitions of validated include “the process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose” [OECD, 2005]; “the process by which the reliability and relevance of a procedure are established for a

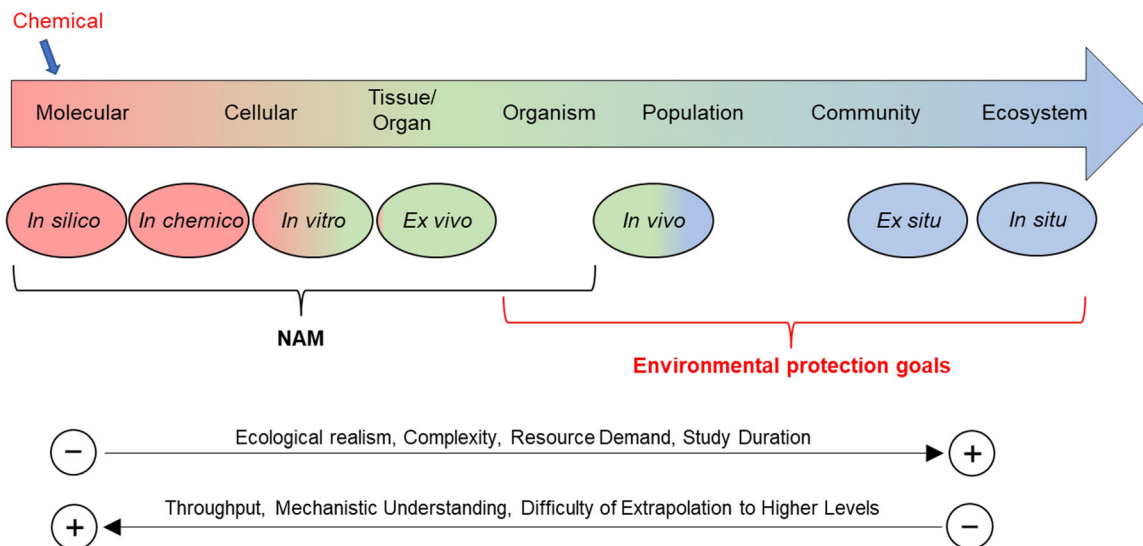


FIGURE 1: Current overview: levels of biological organization and tools in support of chemical ecological assessment. NAM = new approach methodology.

specific purpose” [Balls et al., 1995]; and “the process by which the reliability and relevance of a procedure for a specific purpose are established” [National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods & Interagency Coordinating Committee on the Validation of Alternative Methods, 2003]), which has its own set of challenges. For example, validation studies are usually conducted using reference test substances, which typically elicit strong negative or positive responses. However, most “real-world chemicals” tested for regulatory purposes will typically not elicit such clear effects. This has been recognized, with some OECD validation studies now including weakly acting reference chemicals to ensure that assays can cover a realistic spectrum of effect magnitudes. The USEPA is also developing and utilizing reference chemical sets with a broad spectrum of potency for validation purposes.

Combining data from multiple sources, such as via IATAs which use new and existing data in a flexible manner to address regulatory needs, will build confidence in NAMs. Integrated approaches to testing and assessments are widely recognized as the best regulatory practice for the application of NAMs. However, the current OECD process for IATA adoption is rigid, conservative, and lengthy (5–10 years). It typically requires that data are generated using official test guideline methods, which can take many years to obtain. Although some performance-based test guidelines are considered, there need to be faster ways to enable adoption of NAMs more broadly within regulatory agencies. Figure 2 provides a conceptual diagram regarding the integration of NAMs into IATAs.

CURRENTLY AVAILABLE NAMs TO ELUCIDATE ED PATHWAYS AND MECHANISMS

This section highlights the currently available NAM tools (Figure 3): *in silico* approaches, *in vitro* assays, and eleutheroembryo-based assays. Details on each assay are

available in the Supporting tables. Table 1 also demonstrates the regulatory contexts where NAMs are already being applied.

In silico tools

In silico (i.e., computational) tools represent a growing body of NAMs with the potential to accelerate the toxicological screening process. Advantages of these methods are that they often have high speed and throughput, are low-cost, and can assess a large number of molecules with relative ease. Currently, these tools can be used as an alternative or as a supplement to *in vitro* and *in vivo* testing. Most of the available *in silico* models can be described as ligand-based methods or target structure-based approaches. Quantitative structure-activity relationship (QSARs) models are ligand-based methods that estimate activity based on fingerprints and molecular descriptors (Schneider et al., 2019). These tools look for structural alerts in either a particular portion of the molecule or the molecule as a whole. Examples include the OECD QSAR toolbox (qsartoolbox.org), VEGA HUB (www.vegahub.eu), or CAESAR (www.caesar-project.eu). There are QSARs for specific modalities for the endocrine system. Specifically, estrogen receptor (ER), AR, and aromatase inhibition for steroidogenesis have been well established because of interest and regulatory requirements; but models for other pathways, such as thyroid, still require significant development.

Two such QSAR models were developed by consortia of scientists from 35 international organizations led by the USEPA. The Collaborative Estrogen Receptor Activity Prediction Project evaluates chemicals for potential ER activity (Mansouri et al., 2016), and the Collaborative Modeling Project for Androgen Receptor activity (Mansouri et al., 2020) evaluates potential AR activities. Expert modelers and computational toxicology scientists used the data for approximately 1800 chemicals from the ER (Judson et al., 2015) and AR (Kleinstreuer et al., 2017) pathway models (see below, *In Vitro Tools*) to train QSAR models to predict ER and AR activity, respectively. Both of these QSARs are available through the

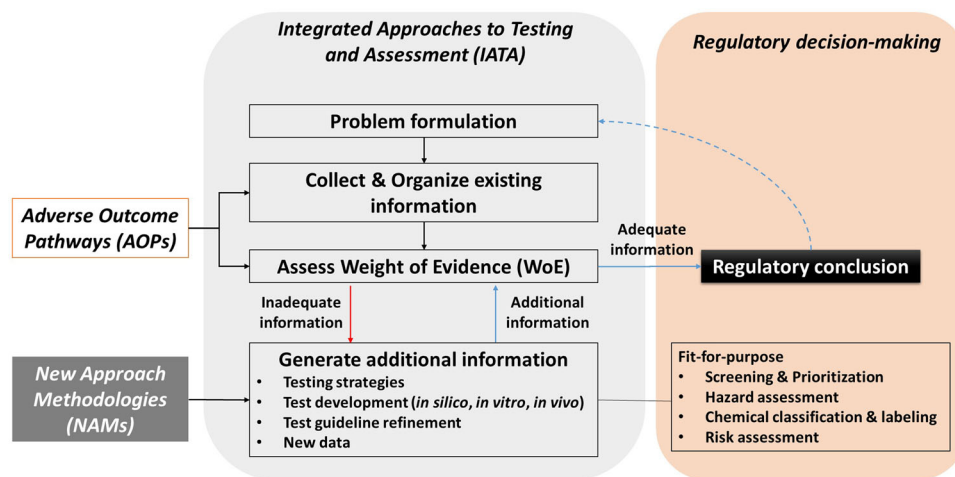
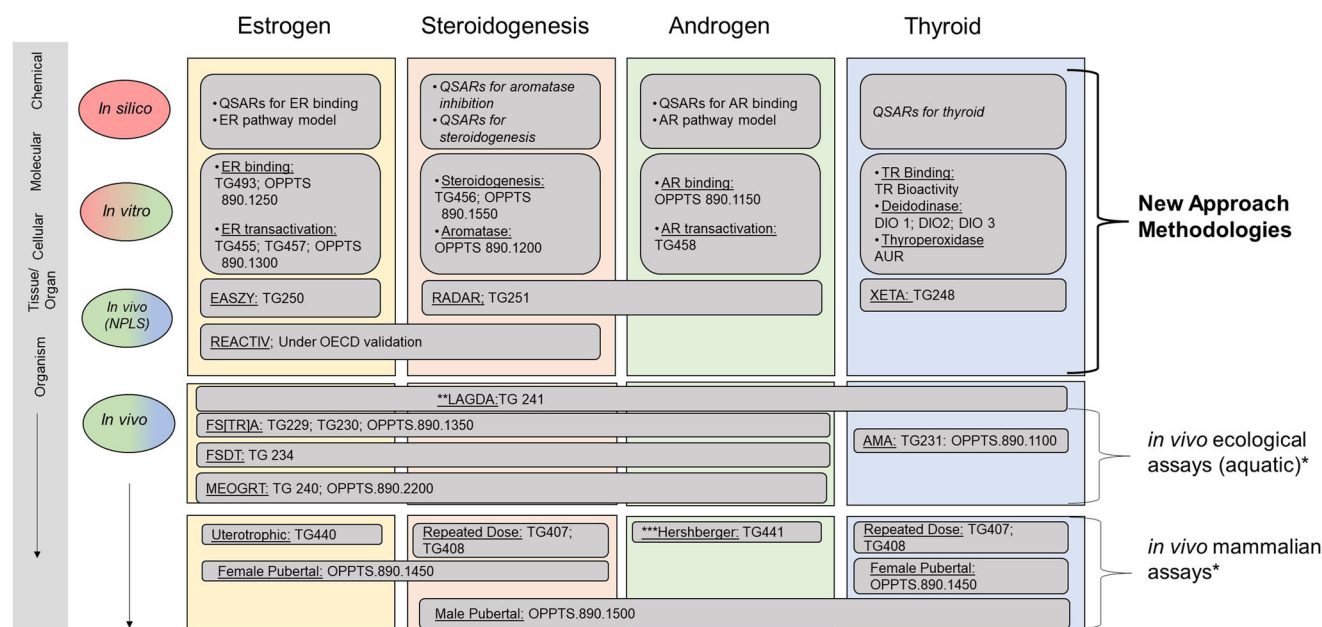


FIGURE 2: An illustration of how new approach methodologies can be integrated to support integrated approaches to testing and assessments. Modified from Organisation for Economic Co-operation and Development (2016). AOP = adverse outcome pathway; NAMs = new approach methodologies; IATAs = integrated approaches to testing and assessments; WoE = weight of evidence.



Italics = in development
 NPLS = non-protected life stages
 *Representative assays
 **Note that, under the current EDSP, the LAGDA is considered primarily for identifying interactions with the hypothalamic–pituitary–thyroid axis.
 ***Hershberger assay specifically measures 5 α -reductase inhibition as well

FIGURE 3: New approach methodologies and *in vivo* assays available to evaluate endocrine modalities. Italics indicate in development. Test guidelines: 248 (OECD, 2019), 250 (OECD, 2021b), 251 (OECD, 2022), 407 (OECD, 2008), 408 (OECD, 2018b), 440 (OECD, 2007), 441 (OECD, 2009c), 455 (OECD, 2021c), 456 (OECD, 2011a), 457 (OECD, 2012b), 458 (OECD, 2020), 407 (OECD, 2008), 493 (OECD, 2015c). *Representative assays. **Note that, under the current Endocrine Disruptor Screening Program, the LAGDA is considered primarily for identifying interactions with the hypothalamic–pituitary–thyroid axis. ***Hershberger assay specifically measures 5 α -reductase inhibition as well. NPLS = nonprotected life stage; QSAR = quantitative structure–activity relationship; ER = estrogen receptor; AR = androgen receptor; TR = thyroid hormone receptor; TG = test guideline; OPPTS = Office of Prevention, Pesticides and Toxic Substances; DIO1, DIO2, and DIO3 = iodothyronine deiodinases 1, 2, and 3; AUR = Amplex UltraRed; EASZY = endocrine active substances, acting through estrogen receptors, using transgenic tg(cyp19a1b:GFP) zebrafish embryos; RADAR = rapid androgen disruption activity reporter; XETA = *Xenopus* eleutheroembryonic thyroid assay; REACTIV = rapid estrogen activity tests *in vivo*; OECD = Organisation for Economic Co-operation and Development; LAGDA = larval amphibian growth and development assay; FS[TR]A = fish short-term reproduction assay; FSDT = fish sexual development test; AMA = amphibian metamorphosis assay; MEOGRT = medaka extended one generation reproduction test.

Open Structure–activity/property Relationship App (National Toxicology Program, 2021). In addition, the USEPA’s CompTox Dashboard (Williams et al., 2017) has a “Predictions” tool (USEPA, 2022a) which allows input of a chemical structure to predict a number of toxicological endpoints including ER relative-binding affinity and ER binding.

There are also target structure-based approaches for EDCs. These investigate EDC targets using docking or more computationally intensive approaches such as molecular dynamics. Docking tools are commonly used for virtual screens for thousands of chemicals, as employed in drug discovery. There are many programs available to understand molecular docking and dynamics: A full review of these is beyond the scope of the present study (see Schneider et al., 2019, for a comprehensive summary).

While the above-described computational tools have great utility for screening and prioritization of testing efforts, most have some common limitations. Typically, the chemical space evaluated by these tools is dependent on the amount of high-quality data that are available to train the models, which is often limited. Many models are also dependent on experimental data which, in itself, can contain error and bias, in some cases leading the

models to inaccurate predictions. Currently, many types of chemistries are incompatible for these models, including halogens, metals, and polymers. Chemical mixtures also prove challenging, with models overlooking potential additive or synergistic effects. From an ecotoxicological perspective, many of the existing tools are based on human data, with unknown specificity for fish or amphibians. The ADME properties are also not always incorporated into models and need to be considered (though physiologically based pharmacokinetic modeling could aid with this; Schneckener et al., 2020). Further, most computational models can only ever provide mechanistic information on the lower levels of organization in an AOP; therefore, more data will be needed to link activity detected *in silico* with an adverse outcome at higher levels.

The tools described above use toxicity data and/or chemical structures as their inputs. Other tools evaluate conservation where there is knowledge of targets, such as the USEPA’s Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS; <https://seqapass.epa.gov/seqapass/>) tool. This tool provides a screening approach that allows for extrapolation of toxicity information across hundreds of species. For some model species (e.g., humans, mice, rats, and zebrafish) the

USEPA has large amounts of data regarding their toxicological susceptibility to various chemicals. However, the existing toxicity data for the whole diversity of species are limited. The SeqAPASS tool extrapolates from these data-rich organisms to other unassessed species to predict their specific potential chemical susceptibility. Some chemical classes have relatively well-defined protein targets (e.g., pesticides), and those proteins are curated in the National Center for Biotechnology Information's protein database (www.ncbi.nlm.nih.gov/protein). The SeqAPASS tool evaluates the similarities of protein structure to identify whether a protein target is present for a chemical interaction in other nontarget species. This tool has been applied to endocrine-specific targets, like the ER, the AR, and proteins in the thyroid axis and involved in steroidogenesis (Ankley et al., 2016; LaLone et al., 2018). There are advantages to tools such as SeqAPASS compared with other computational methods. For example, mixtures could be evaluated if their contents are known. Limitations relevant to SeqAPASS are related to the growing sequence information that may have poor annotation or curation and the limitations of sequence data for certain groups of species, particularly invertebrates. In addition, because the chemical's molecular (protein) target is used to query the tool, there needs to be a better understanding (or sharing of knowledge generated in the development of chemicals) as to the chemical–protein interaction that leads to the effectiveness of the chemical.

Overall, computational methods can be used in fit-for-purpose ways as a part of a WoE assessment.

In vitro tools

In vitro tools can provide mechanistic information and evidence on chemical perturbations via molecular initiating events

and KEs in an AOP. In vitro tools can be cell-free or cell-based and, as with other NAMs, are typically used in a WoE approach.

Cell-free assays are available to measure potential endocrine interactions. While there are countless models and targets, in this review we will highlight a few assays that are currently being used or considered for screening associated with chemical safety or risk assessment (Table 1 and Figure 4). When possible, a description of the assay and the associated parameters and performance was generated and is included in the Supporting Information. There are a number of cell-free assays that have been adopted as test guidelines by either the USEPA or the OECD. The USEPA's Endocrine Disruptor Screening Program (EDSP) Tier 1 battery of assays (USEPA, 2022a, 2022b) includes three test guidelines for cell-free assays: an AR-binding assay (rat ventral prostate cytosol, OPPTS 890.1150; USEPA, 2009b), an ER-binding assay (using rat uterine cytosol, OPPTS 890.1250; USEPA, 2009a), and an aromatase inhibition assay (human recombinant, OPPTS 1200; USEPA, 2009c).

There are also several cell-based assays adopted as test guidelines. The USEPA's EDSP Tier 1 battery of assays includes cell-based test guidelines for ER transcriptional activation (OPPTS 890.1300; USEPA, 2009d) and steroidogenesis (OPPTS 890.1550; USEPA, 2009e). The OECD test guidelines include androgen (458, 455, 493; OECD 2020, 2021c, 2015c, respectively) and steroidogenesis (456; OECD, 2011a). There is also a yeast ER assay from the International Organization for Standardization (2018a, 2018b). The above test guidelines that are a part of the USEPA Tier 1 EDSP (see Table 1) are intended to be reviewed as a battery, which includes five in vitro tests and six in vivo tests, to determine potential estrogen, androgen, and thyroid bioactivity.

The Toxicity Forecaster (“ToxCast”) program and the Tox21 (“Toxicity Testing for the 21st Century”) federal consortium

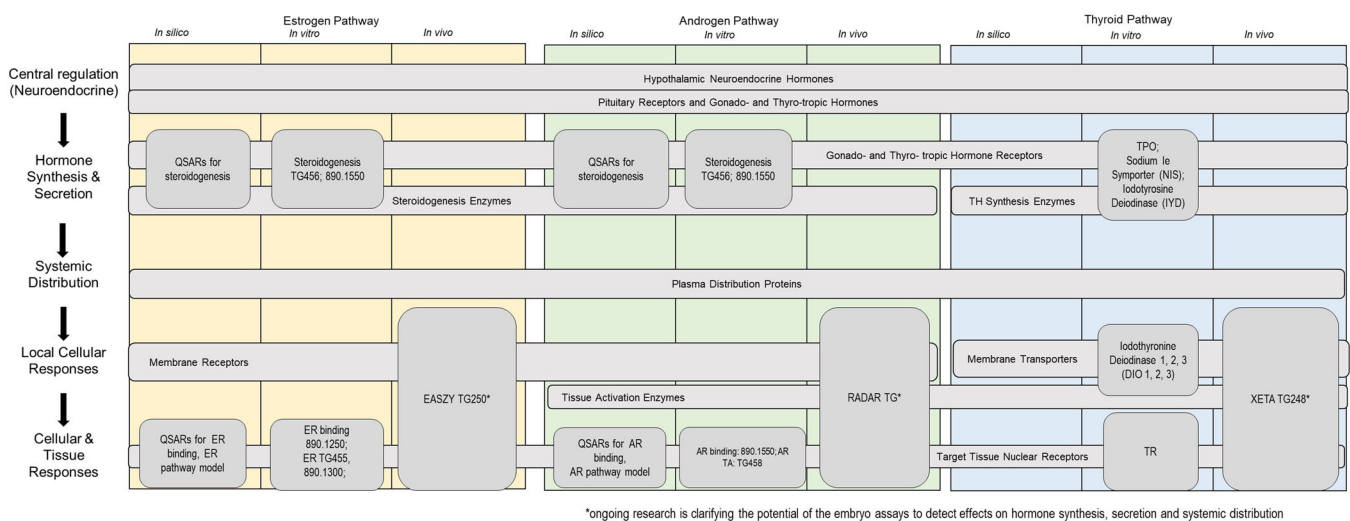


FIGURE 4: Generalized nodes with potential molecular initiating event targets across the estrogen, androgen, and thyroid pathway axes and cross-referenced with available new approach methodologies by type for measuring molecular endpoints. *Ongoing research is clarifying the potential of the embryo assays to detect effects on hormone synthesis, secretion, and systemic distribution. QSAR = quantitative structure–activity relationship; TG = test guideline; TH = thyroid hormone; TPO = thyroperoxidase; NIS = sodium iodide symporter; IYD = iodothyronine deiodinase; ER = estrogen receptor; EASZY = endocrine active substances, acting through estrogen receptors, using transgenic tg(cyp19a1b:GFP) zebrafish embryos; AR = androgen receptor; RADAR = rapid androgen disruption activity reporter; DIO1, DIO2, and DIO3 = iodothyronine deiodinases 1, 2, and 3; TR, thyroid hormone receptor; XETA = *Xenopus* eleutheroembryonic thyroid assay.

prioritize the public release of bioactivity screening data. The USEPA has developed the ToxCast program to make in vitro medium- and high-throughput screening assay data publicly available for the prioritization and hazard characterization of thousands of chemicals. ToxCast contains a variety of assay technologies to evaluate the effects of chemical exposure on diverse biological targets from distinct proteins to more complex cellular processes, including all Tox21 assays. The USEPA's CompTox Chemicals Dashboard (<https://comptox.epa.gov/dashboard>) integrates physicochemical, environmental fate and transport, exposure, usage, in vivo toxicity, and in vitro bioassay data (i.e., ToxCast) for over 870 000 chemicals (Williams et al., 2017). A subset of ToxCast methods and data from this effort is directly relevant to screening for endocrine activity. For example, ToxCast has 18 high-throughput assays based on the molecular events associated with ER activation (e.g., binding, dimerization, transcription, and translation), and the data from these assays were assessed for 1800 chemicals to build a statistical model called the *ER pathway model* (Judson et al., 2015). Similarly, ToxCast has 14 high-throughput assays based on the molecular events associated with AR activation, and the data from 11 assays were assessed for the same set of 1800 chemicals to build the *AR pathway model* (Judson et al., 2020; Kleinstreuer et al., 2017). These models can also provide a measure of potency based on consensus models, where relative potency is scored against the bioactivity of known potent substances (e.g., 17 β -estradiol). The data for each of the ER and AR high-throughput assays and scores for both the ER and AR pathway models are available on the Chemicals Dashboard. In 2015, the USEPA published a *Federal Register* notice seeking comment on using the ER pathway model as a validated alternative for multiple Tier 1 assays (USEPA, 2015a). Very recently, the USEPA published a white paper supporting the validation of these models as alternatives for EDSP Tier 1 screening (USEPA, 2023). The ToxCast ER and AR high-throughput assay data were reassessed to demonstrate that subsets of assays can provide close to the same level of predictivity as the full assay sets listed above for ER agonism (Judson et al., 2017), AR agonism, and AR antagonism (Judson et al., 2020).

The Chemicals Dashboard also contains ToxCast data from assays targeting multiple thyroid hormone receptors (TRs; Paul-Friedman et al., 2019), thyroperoxidase activity (Paul-Friedman et al., 2016), and steroidogenesis (HT-H295R; Haggard et al., 2017). New ToxCast assays for a variety of thyroid pathways are in development. These include the iodothyronine deiodinase 1 (DIO1) inhibition assay (Hornung et al., 2017), the DIO2 inhibition assay (Olker et al., 2018), the DIO3 inhibition assay (Olker et al., 2018), the iodotyrosine deiodinase inhibition assay (Olker et al., 2021), and the sodium iodide symporter inhibition assay (J. Wang et al., 2019).

The currently available in vitro tools have common strengths in that they are quick to perform, many have been adapted to be high-throughput, and they provide specific mechanistic information on potential endocrine activity. The cell-based in vitro assays also share some common strengths, utilizing immortalized and commercially available cell lines, which are

easier to acquire and handle than primary cells and do not need to be sourced from animals. They also can give some idea of potency based on response or binding (Asnake et al., 2019). However, the immortalized cell lines used in the above assays are almost exclusively derived from tumors and are prone to the *Warburg effect* (increase in the rate of glucose uptake and preferential production of lactate, even in the presence of oxygen; Oliveira et al., 2015), which could be a complication for predicting responses to endogenous cells, receptors, and enzymes. Not all modalities and mechanisms are fully covered by existing assays, and there are few that are specific for fish or amphibians. Some cell-free assays require isolation of biological material from animal organs (e.g., the AR-binding assay uses rat ventral prostate cytosol). Using biological material derived from animals increases cost and effort; however, those organs have endogenous expression of a mix of receptors (e.g., uterus expresses ER α and ER β) that may not be found in immortalized cell lines. In addition, receptor-binding assays cannot distinguish if a chemical is an agonist or antagonist, and most in vitro assays would likely not be able to capture responses from life stages when an organism would be more susceptible to the effects of an EDC.

Eleutheroembryo assays

Eleutheroembryo assays have been developed in amphibians and fish with the specific purpose of detecting endocrine activity. The amphibian *Xenopus laevis* is used in the *Xenopus* eleutheroembryonic thyroid assay (XETA; OECD test guideline 248; OECD, 2019). Medaka (*Oryzias latipes*) and zebrafish (*Danio rerio*) are fish species for which eleutheroembryo assays have been developed to investigate estrogenic (endocrine active substances, acting through estrogen receptors, using transgenic tg[cyp19a1b:GFP] zebrafish embryos; OECD test guideline 250 zebrafish eleutheroembryos; OECD, 2021b; rapid estrogen activity tests in vivo [REACTIV] assay, medaka eleutheroembryos) and androgenic (rapid androgen disruption activity reporter [RADAR] assay; OECD test guideline 251 medaka eleutheroembryos; OECD, 2022) activities. Both the REACTIV and RADAR assays also cover downstream steps (5 α -reductase and aromatase) of steroidogenesis. These assays use organisms in the eleutheroembryonic stage that are not capable of independent feeding. At this stage the eleutheroembryos are not considered capable of experiencing pain, distress, suffering, or lasting harm (European Food Safety Authority [EFSA], 2005). However, it is important to note that these eleutheroembryo assays do not necessarily use fewer individuals compared with other OECD Level 3 mechanistic assays (e.g., test guidelines 231, 229; OECD, 2009b, 2012a, respectively).

Currently, the available eleutheroembryo assays do not allow for the identification of specific endocrine targets because they are whole organisms which may include multiple mechanisms by which an endocrine active substance can act. By design, the assays are based on the use of genetic constructs that combine the promoter of an endocrine receptor

and the reporter gene of a green fluorescent protein (GFP). The assays measure the capacity of a substance to activate or inhibit the transcription of the genetic construct, which can happen via different mechanisms. For example, the XETA uses the promoter of the thyroid hormone/basic leucine zipper (TH/bZIP) gene, the expression of which is regulated directly at the onset of metamorphosis by TH (Furlow & Brown, 1999). The transcription of TH/bZIP-GFP is regulated either directly through binding to the TR or modifying the binding of thyroid hormones or indirectly by modifying the amount of hormone available to activate the receptor. Similarly, the REACTIV and RADAR assays can detect substances that directly bind the ER or AR (i.e., ER and AR agonists), respectively, and substances that affect the enzymes of the downstream steps of steroidogenesis (i.e., 5 α -reductase and aromatase). It is acknowledged that organisms in the eleutheroembryonic stage do not produce (sufficient) endogenous hormones like thyroid hormones or steroids. However, recent studies with fish embryos to assess the effects of the potent estrogen 17 α -ethynylestradiol have shown that nonregulated early life stages respond in a manner and with a sensitivity that are similar to sexually mature fish (i.e., production of vitellogenin; Alcaraz et al., 2021). The detection of antagonistic effects, therefore, requires additional treatments where an exogenous hormone is coexposed in the test design.

In the absence of *in vitro* data, the eleutheroembryo assays can only inform on the potential for a test substance to interact directly or indirectly with an endocrine axis via various mechanisms. The level of specificity toward the specific endocrine mechanisms is, however, greater for the eleutheroembryo assays than for their *in vivo* screens (amphibian metamorphosis assay [AMA]; OECD test guideline 231; OECD, 2009b) or the fish short-term reproduction assay (FSTRA; OECD test guideline 229; OECD, 2012a), where nonendocrine mechanisms can also induce the response of most (if not all) measured apical endpoints (i.e., development or reproduction).

There are still uncertainties concerning the metabolic capabilities of eleutheroembryos. The early developmental stages covered have metabolic enzyme expression that is different from that of more developed and adult individuals. Current limitations to the use of eleutheroembryonic assays in regulatory endocrine assessment are mainly related to establishing applicability domains for the mechanisms covered and capacity and experience in laboratories, combined with limited regulatory acceptance. European regulatory authorities are becoming more familiar with the use of the XETA in EDC assessments. The European Chemicals Agency (ECHA) and EFSA recently amended the guidance for the identification of endocrine disruptors in the context of Regulations (EU) No. 528/2012 and (EC) No. 1107/2009, with an annex describing the conditions of use of the XETA as a mechanistic assay to detect thyroid active substances (Arena & van der Linden, 2021). In this document, the XETA is recommended as an alternative to the AMA under certain conditions. The potential of the XETA to predict the outcome of the AMA has been recently investigated using 22 chemicals that have been tested in both assays, showing a high concordance between the outcomes of

the two assays for thyroid active and inactive chemicals (D. Du Pasquier, Laboratoire Watchfrog, F-91000 Evry, France, unpublished data). The forthcoming publication of fish eleutheroembryo assays as OECD test guidelines will further contribute to promoting the regulatory use of this type of assay as NAMs for assessment of endocrine activity.

INTEGRATION AND APPLICATION OF NAMs

NAMs and AOPs

Despite some of the limitations related to NAMs described above, these tools have great promise for integration into the AOP framework. Within the AOP framework, NAMs primarily fit to the “left side” (Figure 1), providing endocrine activity information at the molecular, cellular, and tissue levels. Eleutheroembryo assays can provide information on the whole organism level, although they may be limited by developmental stage and metabolic competence. Figure 4 shows the generalized nodes, neuroendocrine signaling, hormone synthesis, systemic distribution, and cellular and tissue responses, of the endocrine pathways for the estrogen, androgen, and thyroid hormone axes. Within each of these nodes, there are different mechanisms which could be targets for chemical perturbation. The figure shows the NAMs currently available or in development at each of the pathway nodes and can be used to help identify gaps in the biological coverage and potentially inform future NAM development, especially for the assessment of thyroid activity.

There are currently 20 endocrine AOPs registered at the OECD, of which three are endorsed, seven are under review, and 10 are under active development (Table 2). These AOPs cover the estrogen, androgen, thyroid, and steroidogenesis modalities in mammals and fish. Various NAMs are deployed as the detection/quantification methods for measuring the molecular initiating events or early KEs in these AOPs (Table 2). Many NAMs can generate temporal and dose–response data (e.g., the US Tox21 quantitative high-throughput screening program), not only providing empirical support for WoE assessment of the AOPs but also forming the basis for the construction of quantitative prediction models (i.e., quantitative AOPs) in the future. Conversely, connecting to AOPs may facilitate the development of NAMs toward regulatory acceptance because AOPs serve as bridges to link the molecular/cellular effects measured by NAMs to adverse effects of regulatory concern. However, a current weakness is the difficulty in connecting these molecular initiating events and KEs to adverse effects at the organism and population levels aligned to the protection goal of most environmental assessments (Lagadic et al., 2020).

To address this challenge, putative AOPs may be developed where molecular/cellular responses to a chemical exposure are experimentally determined using *in silico* and *in vitro* assays, which are then linked to KEs leading to an adverse outcome. For the environmental endocrine assessment, these adverse outcomes will need to be empirically demonstrated as directly

TABLE 2: An overview of the Organisation for Economic Co-operation and Development adverse outcome pathways on vertebrate endocrine disruption

OECD project	Status	ED modality	AOP title	Type of NAM deployed
1.12	EAGMST under review	Estrogen	Estrogen receptor antagonism leading to reproductive dysfunction	In vitro, in vivo
1.73	Under development	Estrogen	Binding to ER- α in immune cells leading to exacerbation of systemic lupus erythematosus	In vitro, in vivo
1.79	Under development	Estrogen	Early-life stromal ER activation by ED chemicals in the mammary gland leading to enhanced cancer risk	Not described
1.12	TFHAWNT endorsed	Steroidogenesis	Aromatase inhibition leading to reproductive dysfunction	In vitro, in vivo
1.21	EAGMST under review	Steroidogenesis	Aromatase (Cyp19a1) reduction leading to impaired fertility in adult female	In vitro, in vivo
1.9	Under development	Steroidogenesis	5 α -reductase inhibition leading to short anogenital distance in male (mammalian) offspring	In vitro, in vivo
1.9	Under development	Steroidogenesis	Decreased testosterone synthesis leading to short anogenital distance in male (mammalian) offspring	In vitro, in vivo
1.12	TFHAWNT endorsed	Androgen	Androgen receptor agonism leading to reproductive dysfunction (in repeat-spawning fish)	In vitro, in vivo
1.26	Under development	Androgen	Androgen receptor activation leading to hepatocellular adenomas and carcinomas (in mouse and rat)	In vitro
1.9	Under development	Androgen	Androgen receptor antagonism leading to short anogenital distance in male (mammalian) offspring	In vitro
1.1	TFHAWNT endorsed	Thyroid	Inhibition of thyroperoxidase and subsequent adverse neurodevelopmental outcomes in mammals	In vitro, in vivo
1.35	EAGMST under review	Thyroid	Deiodinase 1 inhibition leading to increased mortality via reduced anterior swim bladder inflation	In vitro, in vivo
1.35	EAGMST under review	Thyroid	Deiodinase 1 inhibition leading to increased mortality via reduced posterior swim bladder inflation	In vitro, in vivo
1.35	EAGMST under review	Thyroid	Deiodinase 2 inhibition leading to increased mortality via reduced anterior swim bladder inflation	In vitro, in vivo
1.35	EAGMST under review	Thyroid	Deiodinase 2 inhibition leading to increased mortality via reduced posterior swim bladder inflation	In vitro, in vivo
1.35	EAGMST under review	Thyroid	Thyroperoxidase inhibition leading to increased mortality via reduced anterior swim bladder inflation	In vitro, in vivo
1.59	Under development	Thyroid	Inhibition of thyroid peroxidase leading to impaired fertility in fish	In vitro, in vivo
1.9	Under development	Thyroid	Upregulation of thyroid hormone catabolism via activation of hepatic nuclear receptors and subsequent adverse neurodevelopmental outcomes in mammals	Not described
1.4	Under development	Thyroid	Kidney dysfunction by decreased thyroid hormone	In vitro, in vivo
1.84	Under development	Thyroid	Thyroid receptor antagonism and subsequent adverse neurodevelopmental outcomes in mammals	In vivo

OECD = Organisation for Economic Co-operation and Development; ED = endocrine disruption/disruptor; AOP = adverse outcome pathway; EAGMST = Extended Advisory Group on Molecular Screening and Toxicogenomics; ER = estrogen receptor; TFHA = Task Force on Hazard Assessment; WNT = Working Group of the National Coordinators for the Test Guidelines Programme; Cyp19a1 = cytochrome P450 19a1.

relevant to the population, or the AOP will need to incorporate an additional modeling step to estimate if the individual-level effects translate to the apical assessment endpoints at the population level. Once constructed, these AOPs may be used qualitatively to evaluate chemicals for potential toxicity and, in some cases, quantitatively to build mechanistic models for predicting dose–response relationships and risk estimates (El-Masri et al., 2016). However, the application of AOPs necessitates the extrapolation of in vitro data to in vivo doses, which are converted to external exposure levels where regulatory limits are set. Such in vitro to in vivo extrapolations (IVIVE) help interpret the biological plausibility of in vitro concentrations that induce effects by placing them into the context of tissue concentrations. However, linking tissue concentrations to external doses requires an understanding and characterization of ADME for each chemical (Cohen Hubal et al., 2019; Strikwold et al., 2013). Species differences in ADME also need to be considered (e.g., ADME in humans is not the same as ADME in fish and/or amphibians).

NAMs and IATAs

One barrier to the adoption of some NAMs in certain regulatory contexts is the lack of harmonized, internationally recognized guiding principles on how and when to apply and interpret these new approaches and data for decision-making (Webster et al., 2019). Integrated approaches to testing and assessments are strategies that use new and existing data in a flexible manner to address regulatory needs (Figure 2). To gain acceptance of NAMs, multiple programs, including Accelerating the Pace of Chemical Risk Assessment (Kavlock et al., 2018), EU ToxRisk (Moné et al., 2020), and the OECD IATA Programme, are using case studies to explore how to apply and interpret NAMs in a decision-making context. This should lead to more widespread acceptance of these tools.

New approach methodologies integrated into AOPs, IATAs, and other WoE paradigms hold promise for more efficient and timely chemical assessments based on sound. However, they are not one-to-one replacements for in vivo tests and should be used in a battery and applied in a WoE approach. If incorporated in this way, NAMs could be used to simultaneously improve the science of endocrine activity and disruption assessment by providing a better mechanistic understanding compared with animal data alone.

INCREASING REGULATORY ACCEPTANCE OF NAMs

There has been a consistent stream of NAM development within the fields of both human health and ecological safety assessment over the past decade, sparked by efforts to reduce the need for more resource-intensive animal testing. In turn, the shift toward a more mechanistic focus and the increase in the use and application of AOPs have provided the platform and framework to support the further development, interpretation, and use of NAMs. Particularly in the endocrine

space, robust knowledge of the estrogen, androgen, thyroid, and steroidogenesis pathways and a high degree of conservation across species afford a unique opportunity to utilize and apply NAMs within regulatory contexts but also in ecological risk assessments more broadly. While not exhaustive, several short- and longer-term priorities to help facilitate the regulatory use of NAMs are discussed below.

Short-term

- **Define and increase the applicability domains of available NAMs:** For assays already well developed and in use, approaches to expand their use and establish agreed chemical and taxonomic applicability domains should be investigated so that information can be incorporated into regulatory guidance. For example, greater application of NAMs could be encouraged by use of performance-based testing guidelines. Greater latitude in the experimental design is allowed if certain “performance” criteria are still met. This means that laboratories with analogous methods can still apply their models within a test guideline framework that in turn should be more broadly acceptable to regulatory authorities and study sponsors. Recent experience has also demonstrated the need for greater clarity around applicability domains under which the data can be considered reliable for regulatory decision-making. One recent example is the appendix (Arena & van der Linden, 2021) to the ECHA/EFSA guidance document for the application of the XETA test guideline for plant protection products in the European Union. This appendix has increased regulatory clarity for registrants to decide when the XETA assay is appropriate over the performance of the in vivo AMA. This highlights the importance of the international validation program to ensure that there is clarity on the applicability domains reliably covered (or not) as demonstrated in the validation exercise. These activities all work to increase confidence within regulatory authorities, which in turn provides reassurance to registrants that NAM approaches are likely to be accepted in place of in vivo approaches or at least reliably inform when in vivo studies are truly necessary.
- **Further develop NAMs for the thyroid pathway:** The thyroid modality remains an area where, despite the development of numerous NAM assays for molecular initiating events, there are no validated in vitro test guidelines (Noyes et al., 2019) or OECD conceptual framework Level 2 assays. Development of additional NAMs to evaluate thyroid effects is needed. At present, the use of in vitro data beyond screening and prioritization for thyroid modality is challenged by the complexity of potential thyroid-related adverse outcomes and the limited knowledge of mechanistic processes controlling such responses. There are emerging tools like those described above, as well as the Danish QSAR Database which includes thyroid disruption models (Rosenberg et al., 2017). Additional research into the role of the thyroid in fish and amphibians may be needed before NAM tools can be developed (Knäpen et al., 2020).

- **Increase understanding of ADME processes related to NAMs:** Incorporating ADME data with NAMs will increase scientific confidence. Examples of recent work to do this include the USEPA's development of the Alginate Immobilization of Metabolic Enzymes platform to include metabolic competence in an ER transactivation assay (Deisenroth et al., 2020) and additional work that introduces metabolic capacity into the cells by performing transfections with modified mRNAs (DeGroot et al., 2018). Quantitative IVIVE (QIVIVE), which uses concentration responses from in vitro toxicodynamic or toxicity assays along with in vitro toxicokinetic measures (e.g., plasma binding) as inputs to a physiologically based pharmacokinetic model to derive circulating plasma concentrations (Bell et al., 2018; Wetmore, 2015), is being used to predict in vivo toxicity. Quantitative in vitro to in vivo extrapolation is increasingly being used to predict toxicity with human (Beames et al., 2020; Clewell et al., 2019; Paul Friedman et al., 2020; Shi et al., 2020) and rat (Honda et al., 2019; Scholze et al., 2020) models; however, the process has not yet been applied to fish or amphibian endocrine test endpoints. Considerations for implementation of QIVIVE to a fish model include appropriate in vitro toxicity assays for fish, in vitro toxicokinetic assays, and physiologically based pharmacokinetic models capable of calculating external aquatic exposure concentrations from predicted plasma concentrations. Assays for plasma binding (Henneberger et al., 2020) and intrinsic metabolic clearance (Black et al., 2021; Nichols et al., 2018) have been developed for representative fish species, as have detailed models (Brinkmann et al., 2014; Grech et al., 2019; Nichols et al., 2004, 2018; Stadnicka et al., 2012). Many components necessary for the implementation of QIVIVE for fish models are available; however, more toxicokinetic data are needed to develop the requisite data sets for comparison with fish endpoints relevant for screening or testing for ED.
- **Curate high-quality in vivo data in standardized formats:** For the foreseeable future, the performance of NAMs will be benchmarked against the results and decisions that have been made using in vivo study data. Therefore, it is imperative to make high-quality in vivo data accessible for the relevant comparisons to be made in assessing NAM performance. Despite the wealth of in vivo toxicity data that exist, finding, collecting, and assessing these data can be challenging. Many publicly available regulatory submissions would require manual, time-consuming document extraction; and reports may not be made available immediately or may be redacted because of the inclusion of proprietary information. To facilitate the incorporation of NAMs into the EDC regulatory space, effort needs to be made to map what data exist and ensure that the data are available in a useable format. Advances in data curation science and machine learning data tools hold great promise in addressing this challenge. Another difficulty lies in the lack of data standardization that exists across sectors, regions, and individual studies. Looking forward, facilitating data interoperability and usage will require harmonized language and data templates, preferably located within a centralized international

database. A parallel project mentioned previously assessing the robustness of the current in vivo assays has established databases of regulatory data for the AMA, medaka extended one generation reproduction test (OECD test guideline 240; OECD, 2015a) and FSTRA (OECD test guideline 229; OECD, 2012a). These databases could be leveraged to enable the relevant comparisons.

- **Continue to “map” NAMs using the AOP framework:** Identifying where various NAMs fit along an AOP facilitates clear and transparent communication across the scientific and regulatory communities about their use and potential application. In addition, it allows anchoring of various NAMs to adverse outcomes and clearly identifies where there are gaps in method development and/or quantitative understandings across KEs. The continued mapping and usage of the AOP framework should be encouraged.
- **Develop case studies:** Where available, case studies for chemicals and chemical classes with existing NAM and in vivo data should be analyzed to compare the differences in conclusions from assessments. This could support NAMs being used for risk or hazard assessment instead of just prioritization or screening. These case studies could incorporate paradigms like IATAs or AOPs to use multiple evidence streams in a WoE analysis for evaluation. This type of work is being initiated within this collaboration.
- **Consistent dialogue across the scientific and regulatory communities:** Engaging regulatory authorities early on in defining data needs is critical and allows greater provision of opportunities where data generated using NAMs can be submitted and assessed in parallel with traditional in vivo data to build confidence. Such an approach has been encouraged in data requests on substances undergoing evaluation or reevaluation at the Joint FAO/WHO Meeting on Pesticide Residues, with the output offering to “evaluate without prejudice, in parallel, any data generated using emerging methods that in the view of sponsors could substitute for information obtained using conventional testing methods” (Food and Agriculture Organization & WHO, 2021). With time and experience, it could be assessed whether both in vitro and in vivo approaches are always needed to make appropriate decisions. The normalization of including NAM data within submissions would also gain acceptance among regulatory scientists who interpret these different data sources.

Longer-term

- **Continued funding and allocation of resources to support NAMs:** Consistent and adequate allocation of resources is needed to validate and establish which NAMs (or battery of NAMs) are best able to reduce or replace and improve on in vivo testing approaches. Validation work tends to lack specific funding mechanisms and is not viewed as being as “exciting” as basic research, but it is the most critical step toward regulatory use and application. Validation will encourage the use and uptake of these methods as well as opportunities to

incorporate additional endpoints into these assay batteries, which could reduce overall time and costs.

- **Development of quantitative linkages across KEs:** Generally, AOPs are described in qualitative terms: They usually do not provide an understanding of quantitative relationships between the molecular initiating events, subsequent KEs, and the ultimate adverse outcome (Sewell et al., 2018). The important, and not yet met, requirement is that AOPs reflect the thresholds required for progression of a defined biological perturbation to a specific toxicity. In all areas, there may be some data regarding the thresholds characterizing KE relationships that already exist and may include information on dose–response relationships and temporal concordance between KEs. However, often these data are not appropriately captured or made available in a readily accessible form. If supporting evidence for KEs and KE relationships was suitably annotated with dose and time information, then it would be possible for temporal and dose concordance between KEs to be inferred computationally based on Bradford-Hill considerations (Becker et al., 2015). There has been some notable movement in this area to date (Conolly et al., 2017; Villeneuve et al., 2021), but more needs to be done.
- **Further development of new technologies and approaches:** Continued support of methods that are under development and have potential to support quantitative AOP development and application in the ecological space (for ED and otherwise) is needed. A brief description of several technologies that are in development and could be used for endocrine modalities is provided below. This is not meant to be exhaustive.
 - **High-content imaging-based high-throughput phenotypic profiling:** Also known as *cell painting*, high-content imaging-based high-throughput phenotypic profiling is a NAM that has mostly been piloted in human cells (Ljosa et al., 2013). As opposed to targeted in vitro assays, profiling assays collect multiplexed measurements of different features and evaluate the global response of intact cells to chemicals. For toxicology, it is currently being used at the USEPA to compare the results of in vitro studies to values from existing mammalian toxicity studies (Nyffeler et al., 2020). More work needs to be done to gain confidence in this method for human cells before applying to ecotoxicology, specifically for endocrine assessment.
 - **Transcriptomics experiments:** Transcriptomics experiments are routinely run for many chemicals, but there remains difficulty interpreting pathway and/or adverse effects associated with differential gene expression (Pain et al., 2020). Greater investment in sequencing tools and genome annotation could aid the reduction of resource-intensive in vivo studies by in vitro replacements or refinement of experiments, resulting in more mechanistic information. This would also support advances in species extrapolation and better understanding of species differences. Transcriptomic assays provide mechanistic information on the specific KEs triggered by a compound but also reflect a whole-organism response that could ultimately replace currently required confirmatory

higher-tier in vivo testing. While transcriptomics shows great promise and is increasingly included as supplementary information in support of chemical hazard assessment, more work is needed to support the adoption of these tools, in particular for EDC-specific outcomes. Below are some examples:

- In addition to pathway analyses, there is interest in developing transcriptomic-based points of departures. The HESI's Emerging Systems Toxicology for Assessment of Risk Committee is currently exploring the use of transcriptomic points of departures for mammalian studies for chemicals (Johnson et al., 2022). Regarding ecotoxicity assessments, recent research has shown that transcriptomic points of departure from short-term exposure studies are protective of chronic effects for diverse fish species exposed to estrogenic chemicals (Alcaraz et al., 2021; Pagé-Larivière et al., 2019). New, user-friendly, and intuitive tools such as FastBMD (<https://www.fastbmd.ca>) now facilitate the derivation of transcriptomics benchmarks.
- Building on the potential for transcriptomics and other approaches to support cross-species extrapolation, a multisector consortium aims to improve the International Consortium to Advance Cross-Species Extrapolation in Regulation (LaLone et al., 2021).
- There have also been efforts to develop and standardize reduced transcriptome panels that are representative of diverse biological pathways of interest from a toxicological perspective (Soufan et al., 2019; P. Wang et al., 2018). One such example is EcoToxChips, which uses early, nonprotected vertebrate embryonic life stages (birds, amphibians, and fish) in combination with an intuitive and standardized bioinformatics platform (EcoToxXplorer.ca) for chemical prioritization and possible hazard assessment (Basu et al., 2019; Soufan et al., 2019).
- **Organ on a chip:** These assays utilize three-dimensional microfluidic cell culture to simulate an entire organ (Low et al., 2021). These models have been growing in popularity and have been used in the drug discovery process and for toxicology research. For ecotoxicology, there are various models including Fish-on-a-Chip, which uses microfluids to imitate zebrafish-based research (Yang et al., 2016), and Fish-Gut-on-Chip, which reconstructs the fish intestinal barrier (Drieschner et al., 2019). Work needs to be done to apply these tools to endocrine modalities. While these technologies could reduce the need for in vivo testing, limitations include the inability to represent the whole organism organ level because these are only organ-specific functions and do not capture organ-to-organ communication via endocrine pathways. The metabolic ability of these systems varies, as does their ability to capture organ-to-organ communication via endocrine pathways. Multiorgan systems are undergoing evaluation using combinations of up to three organ systems in other research areas however, and this type of approach may support use of organ-on-a-chip for endocrine assessments.

- **Exploration of other pathways and applications:** Endocrine pathways other than estrogen, androgen, thyroid, and steroidogenesis often lack NAMs (e.g., aryl hydrocarbon receptor, peroxisome proliferator-activated receptors, and glucocorticoid signaling; Martyniuk et al., 2021). In general, many of the concepts in the present study could apply to other modalities and other mechanisms of toxicity (e.g., neurotoxicity, genotoxicity) that are being explored via AOPs. Targeted assays could be designed for the modalities and/or combinations with nontargeted techniques could aid in the replacement and reduction of resource-intensive in vivo tests for these pathways. However, it is important to keep in mind the environmental protection goals (e.g., a refined assay for the ER modality that would capture the most important molecular initiating events of the ER axis applicable to population-relevant adverse outcomes). As with all toxicity testing, mixtures are difficult to assess. To deal with mixtures, tools need to be improved or developed to be more representative of actual organisms and environmental conditions, and cumulative exposures need to be considered for endocrine assessment.

CONCLUSIONS

New approach methodologies can be utilized to both reduce the number of animals needed to assess a chemical's endocrine potential and simultaneously improve the science of endocrine activity and disruption assessment. While NAMs are currently used in some regulatory assessments and when data are available, advances need to be made to be able to move away from in vivo testing to a fit-for-purpose test battery of NAMs. New approach methodologies should be integrated as part of a WoE approach for hazard or risk assessment using the AOP framework. The application and use of NAMs within a WoE approach may satisfy outstanding data needs, testing requirements, or protection goals. Immediate next steps to advance adaptation of NAMs in regulatory contexts include performing case studies on compounds that have been evaluated with NAMs and in vivo methods in a regulatory context and comparing the outcomes.

Supporting Information—The Supporting Information is available on the Wiley Online Library at <https://doi.org/10.1002/etc.5584>.

Acknowledgments—This Health and Environmental Sciences Institute (HESI) scientific initiative is primarily supported by in-kind contributions (from public and private sector participants) of time, expertise, and experimental effort. These contributions are supplemented by direct funding (that largely supports program infrastructure and management) that was provided by HESI's corporate sponsors. A list of supporting organizations is available at www.hesiglobal.org. The present study was supported in part by HESI's Next Generation Ecological Risk Assessment Committee. We also thank E. McDermott for her help with formatting and checking the information in the Supporting tables.

Disclaimer—The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The opinions presented are those of the authors and not necessarily their organizations.

Author Contributions Statement—**Constance A. Mitchell:** Conceptualization; Data curation; Funding acquisition; Methodology; Project administration; Resources; Supervision; Visualization; Writing—original draft; Writing—review & editing. **Natalie Burden:** Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Visualization; Writing—original draft; Writing—review & editing. **Mark Bonnell:** Conceptualization; Investigation; Methodology; Resources; Writing—original draft; Writing—review & editing. **Markus Hecker, Carlie A. LaLone:** Conceptualization; Data curation; Writing—original draft; Writing—review & editing. **Tom H. Hutchinson, Magdalena Jagla:** Conceptualization; Data curation; Methodology; Writing—original draft; Writing—review & editing. **Laurent Lagadic:** Data curation; Methodology; Writing—original draft; Writing—review & editing. **Scott G. Lynn:** Conceptualization; Data curation; Investigation; Methodology; Visualization; Writing—original draft; Writing—review & editing. **Bryon Shore:** Data curation; Writing—original draft; Writing—review & editing. **You Song:** Data curation; Methodology; Visualization; Writing—original draft; Writing—review & editing. **Sara M. Vliet:** Data curation; Investigation; Writing—original draft; Writing—review & editing. **James R. Wheeler:** Conceptualization; Data curation; Formal analysis; Methodology; Project administration; Visualization; Writing—original draft; Writing—review & editing. **Michelle R. Embry:** Conceptualization; Methodology; Project administration; Supervision; Writing—original draft; Writing—review & editing.

Data Availability Statement—There is no data generated in this review.

REFERENCES

- Alcaraz, A. J. G., Mikulášek, K., Potěšil, D., Park, B., Shekh, K., Ewald, J., Burbridge, C., Zdráhal, Z., Schneider, D., Xia, J., Crump, D., Basu, N., & Hecker, M. (2021). Assessing the toxicity of 17 α -ethinylestradiol in rainbow trout using a 4-day transcriptomics benchmark dose (BMD) embryo assay. *Environmental Science & Technology*, 55(15), 10608–10618. <https://doi.org/10.1021/acs.est.1c02401>
- Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., Mount, D. R., Nichols, J. W., Russom, C. L., Schmieder, P. K., Serrano, J. A., Tietge, J. E., & Villeneuve, D. L. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry*, 29(3), 730–741. <https://doi.org/10.1002/etc.34>
- Ankley, G. T., LaLone, C. A., Gray, L. E., Villeneuve, D. L., & Hornung, M. W. (2016). Evaluation of the scientific underpinnings for identifying estrogenic chemicals in nonmammalian taxa using mammalian test systems. *Environmental Toxicology and Chemistry*, 35(11), 2806–2816. <https://doi.org/10.1002/etc.3456>
- Arena, M., & van der Linden, S. (2021). Annex A—Use of the XETA in the assessment strategy of the ECHA/EFSA guidance.
- Asnake, S., Modig, C., & Olsson, P.-E. (2019). Species differences in ligand interaction and activation of estrogen receptors in fish and human. *The Journal of Steroid Biochemistry and Molecular Biology*, 195, Article 105450. <https://doi.org/10.1016/j.jsbmb.2019.105450>

- Balls et al. (1995). Validation of alternative tests in the European Union. *Current Problems in Dermatology*, 23, 265–274. <https://doi.org/10.1159/000424322>
- Basu, N., Crump, D., Head, J., Hickey, G., Hogan, N., Maguire, S., Xia, J., & Hecker, M. (2019). EcoToxChip: A next-generation toxicogenomics tool for chemical prioritization and environmental management. *Environmental Toxicology and Chemistry*, 38(2), 279–288. <https://doi.org/10.1002/etc.4309>
- Beames, T., Moreau, M., Roberts, L. A., Mansouri, K., Haider, S., Smeltz, M., Nicolas, C. I., Doheny, D., Phillips, M. B., Yoon, M., Becker, R. A., McMullen, P. D., Andersen, M. E., Clewell, R. A., & Hartman, J. K. (2020). The role of fit-for-purpose assays within tiered testing approaches: A case study evaluating prioritized estrogen-active compounds in an in vitro human uterotrophic assay. *Toxicology and Applied Pharmacology*, 387, Article 114774. <https://doi.org/10.1016/j.taap.2019.114774>
- Becker, R. A., Ankley, G. T., Edwards, S. W., Kennedy, S. W., Linkov, I., Meek, B., Sachana, M., Segner, H., Van Der Burg, B., Villeneuve, D. L., Watanabe, H., & Barton-Maclaren, T. S. (2015). Increasing scientific confidence in adverse outcome pathways: Application of tailored Bradford-Hill considerations for evaluating weight of evidence. *Regulatory Toxicology and Pharmacology*, 72(3), 514–537. <https://doi.org/10.1016/j.yrtph.2015.04.004>
- Bell, S. M., Chang, X., Wambaugh, J. F., Allen, D. G., Bartels, M., Brouwer, K. L. R., Casey, W. M., Choksi, N., Ferguson, S. S., Fraczkiwicz, G., Jarabek, A. M., Ke, A., Lumen, A., Lynn, S. G., Paini, A., Price, P. S., Ring, C., Simon, T. W., Sipes, N. S., ... Kleinstreuer, N. C. (2018). In vitro to in vivo extrapolation for high throughput prioritization and decision making. *Toxicology In Vitro*, 47, 213–227. <https://doi.org/10.1016/j.tiv.2017.11.016>
- Black, S. R., Nichols, J. W., Fay, K. A., Matten, S. R., & Lynn, S. G. (2021). Evaluation and comparison of in vitro intrinsic clearance rates measured using cryopreserved hepatocytes from humans, rats, and rainbow trout. *Toxicology*, 457, Article 152819. <https://doi.org/10.1016/j.tox.2021.152819>
- Brinkmann, M., Eichbaum, K., Buchinger, S., Reifferscheid, G., Bui, T., Schäffer, A., Hollert, H., & Preuss, T. G. (2014). Understanding receptor-mediated effects in rainbow trout: In vitro–in vivo extrapolation using physiologically based toxicokinetic models. *Environmental Science & Technology*, 48(6), 3303–3309. <https://doi.org/10.1021/es4053208>
- Burden, N., Embry, M. R., Hutchinson, T. H., Lynn, S. G., Maynard, S. K., Mitchell, C. A., Pellizzato, F., Sewell, F., Thorpe, K. L., Weltje, L., & Wheeler, J. R. (2022). Investigating endocrine-disrupting properties of chemicals in fish and amphibians: Opportunities to apply the 3Rs. *Integrated Environmental Assessment and Management*, 18(2), 442–458. <https://doi.org/10.1002/ieam.4497>
- Clewell, H. J., Campbell, J. L., Van Landingham, C., Franzen, A., Yoon, M., Dodd, D. E., Andersen, M. E., & Gentry, P. R. (2019). Incorporation of in vitro metabolism data and physiologically based pharmacokinetic modeling in a risk assessment for chloroprene. *Inhalation Toxicology*, 31(13–14), 468–483. <https://doi.org/10.1080/08958378.2020.1715513>
- Cohen Hubal, E. A., Wetmore, B. A., Wambaugh, J. F., El-Masri, H., Sobus, J. R., & Bahadori, T. (2019). Advancing internal exposure and physiologically-based toxicokinetic modeling for 21st-century risk assessments. *Journal of Exposure Science & Environmental Epidemiology*, 29(1), 11–20. <https://doi.org/10.1038/s41370-018-0046-9>
- Collier, Z. A., Gust, K. A., Gonzalez-Morales, B., Gong, P., Wilbanks, M. S., Linkov, I., & Perkins, E. J. (2016). A weight of evidence assessment approach for adverse outcome pathways. *Regulatory Toxicology and Pharmacology*, 75, 46–57. <https://doi.org/10.1016/j.yrtph.2015.12.014>
- Conolly, R. B., Ankley, G. T., Cheng, W., Mayo, M. L., Miller, D. H., Perkins, E. J., Villeneuve, D. L., & Watanabe, K. H. (2017). Quantitative adverse outcome pathways and their application to predictive toxicology. *Environmental Science & Technology*, 51(8), 4661–4672. <https://doi.org/10.1021/acs.est.6b06230>
- Crane, M., Hallmark, N., Lagadic, L., Ott, K., Pickford, D., Preuss, T., Thompson, H., Thorbek, P., Weltje, L., & Wheeler, J. R. (2019). Assessing the population relevance of endocrine-disrupting effects for nontarget vertebrates exposed to plant protection products. *Integrated Environmental Assessment and Management*, 15(2), 278–291. <https://doi.org/10.1002/ieam.4113>
- DeGroot, D. E., Swank, A., Thomas, R. S., Strynar, M., Lee, M. Y., Carmichael, P. L., & Simmons, S. O. (2018). mRNA transfection retrofits cell-based assays with xenobiotic metabolism. *Journal of Pharmacological and Toxicological Methods*, 92, 77–94. <https://doi.org/10.1016/j.vascn.2018.03.002>
- Deisenroth, C., DeGroot, D. E., Zurlinden, T., Eicher, A., McCord, J., Lee, M. Y., Carmichael, P., & Thomas, R. S. (2020). The alginate immobilization of metabolic enzymes platform retrofits an estrogen receptor transactivation assay with metabolic competence. *Toxicological Sciences*, 178(2), 281–301. <https://doi.org/10.1093/toxsci/kfaa147>
- Drieschner, C., Könemann, S., Renaud, P., & Schirmer, K. (2019). Fish-gut-on-chip: Development of a microfluidic bioreactor to study the role of the fish intestine in vitro. *Lab on a Chip*, 19(19), 3268–3276. <https://doi.org/10.1039/C9LC000415G>
- El-Masri, H., Kleinstreuer, N., Hines, R. N., Adams, L., Tal, T., Isaacs, K., Wetmore, B. A., & Tan, Y.-M. (2016). Integration of life-stage physiologically based pharmacokinetic models with adverse outcome pathways and environmental exposure models to screen for environmental hazards. *Environmental Sciences*, 152(1), 230–243. <https://doi.org/10.1093/toxsci/kfw082>
- Embry, M. R., Belanger, S. E., Braunbeck, T. A., Galay-Burgos, M., Halder, M., Hinton, D. E., Léonard, M. A., Lillicrap, A., Norberg-King, T., & Whale, G. (2010). The fish embryo toxicity test as an animal alternative method in hazard and risk assessment and scientific research. *Aquatic Toxicology*, 97(2), 79–87. <https://doi.org/10.1016/j.aquatox.2009.12.008>
- European Chemicals Agency. (2016). *How to use alternatives to animal testing to fulfil the information requirements for REACH registration. Practical guide*. Publications Office of the European Union. <https://doi.org/10.2823/194297>
- European Food Safety Authority. (2005). *Opinion of the scientific panel on animal health and welfare (AHAW) on a request from the Commission related to the aspects of the biology and welfare of animals used for experimental and other scientific purposes*.
- European Chemicals Agency (ECHA) & European Food Safety Authority (EFSA). (2018). Guidance for the identification of endocrine disruptors in the context of Regulations (EU) no 528/2012 and (EC) no 1107/2009. *EFSA Journal*, 16(6), 5311. <https://doi.org/10.2903/j.efsa.2018.5311>
- Food and Agriculture Organization & World Health Organization. (2021, September 6–17 and October 4 and 7). *List of substances scheduled for follow up evaluation or new uses and request for data*. Joint FAO/WHO Meeting on Pesticide Residues 2021 meeting. Online.
- Food and Drug Administration. (1998). *Guidance for industry: Environmental assessment of human drug and biologics applications. Center for drug evaluation and research (CDER) and center for biologics evaluation and research (CBER)*. Silver Spring.
- Food and Drug Administration. (2016). *Guidance document: Environmental assessment: Questions and answers regarding drugs with estrogenic, androgenic, or thyroid activity. Center for drug evaluation and research (CDER)*. Silver Spring.
- Frank R. Lautenberg Chemical Safety for the 21st century Act. Pub. L. No. 114-182, 130 Stat. 448 (2016). <https://www.congress.gov/114/plaws/publ182/PLAW-114publ182.pdf>
- Furlow, J. D., & Brown, D. D. (1999). In vitro and in vivo analysis of the regulation of a transcription factor gene by thyroid hormone during *Xenopus laevis* metamorphosis. *Molecular Endocrinology*, 13(12), 2076–2089. <https://doi.org/10.1210/mend.13.12.0383>
- Grech, A., Tebby, C., Brochot, C., Bois, F. Y., Bado-Nilles, A., Dorne, J.-L., Quignot, N., & Beaudouin, R. (2019). Generic physiologically-based toxicokinetic modelling for fish: Integration of environmental factors and species variability. *Science of the Total Environment*, 651, 516–531. <https://doi.org/10.1016/j.scitotenv.2018.09.163>
- Haggard, D. E., Karmaus, A. L., Martin, M. T., Judson, R. S., Setzer, R. W., & Paul Friedman, K. (2017). High-throughput H295R steroidogenesis assay: Utility as an alternative and a statistical approach to characterize effects on steroidogenesis. *Toxicological Sciences*, 162(2), 509–534. <https://doi.org/10.1093/toxsci/kfx274>
- Henneberger, L., Klüver, N., Mühlenbrink, M., & Escher, B. (2020). Trout and human plasma protein binding of selected pharmaceuticals informs the fish plasma model. *Environmental Toxicology and Chemistry*, 41(3), 559–568. <https://doi.org/10.1002/etc.4934>
- Honda, G. S., Pearce, R. G., Pham, L. L., Setzer, R. W., Wetmore, B. A., Sipes, N. S., Gilbert, J., Franz, B., Thomas, R. S., & Wambaugh, J. F. (2019). Using the concordance of in vitro and in vivo data to evaluate extrapolation assumptions. *PLOS ONE*, 14(5), Article e0217564. <https://doi.org/10.1371/journal.pone.0217564>
- Hornung, M. W., Korte, J. J., Olker, J. H., Denny, J. S., Knutsen, C., Hartig, P. C., Cardon, M. C., & Degitz, S. J. (2017). Screening the ToxCast Phase 1 chemical library for inhibition of deiodinase type 1 activity. *Toxicological Sciences*, 162(2), 570–581. <https://doi.org/10.1093/toxsci/kfx279>

- International Organization for Standardization. (2018a). *Water quality—Determination of the estrogenic potential of water and waste water—Part 1: Yeast estrogen screen (Saccharomyces cerevisiae)* (ISO 19040-1:2018).
- International Organization for Standardization. (2018b). *Water quality—Determination of the estrogenic potential of water and waste water—Part 2: Yeast estrogen screen (A-YES) Arxula adenivorans* (ISO 19040-2:2018).
- International Programme on Chemical Safety. (2002). *Global assessment of the state-of-the-science of endocrine disruptors*. World Health Organization.
- Johnson, K. J., Auerbach, S. S., Stevens, T., Barton-Maclaren, T. S., Costa, E., Currie, R. A., Dalmás Wilk, D., Haq, S., Rager, J. E., Reardon, A. J. F., Wehmas, L., Williams, A., O'Brien, J., Yauk, C., LaRocca, J. L., & Pettit, S. (2022). A transformative vision for an omics-based regulatory chemical testing paradigm. *Toxicological Sciences*, 190, 127–132. <https://doi.org/10.1093/toxsci/kfac097>
- Judson, R., Houck, K., Paul Friedman, K., Brown, J., Browne, P., Johnston, P. A., Close, D. A., Mansouri, K., & Kleinstreuer, N. (2020). Selecting a minimal set of androgen receptor assays for screening chemicals. *Regulatory Toxicology and Pharmacology*, 117, Article 104764. <https://doi.org/10.1016/j.yrtph.2020.104764>
- Judson, R. S., Houck, K. A., Watt, E. D., & Thomas, R. S. (2017). On selecting a minimal set of in vitro assays to reliably determine estrogen agonist activity. *Regulatory Toxicology and Pharmacology*, 91, 39–49. <https://doi.org/10.1016/j.yrtph.2017.09.022>
- Judson, R. S., Magpantay, F. M., Chickarmane, V., Haskell, C., Tania, N., Taylor, J., Xia, M., Huang, R., Rotroff, D. M., Filer, D. L., Houck, K. A., Martin, M. T., Sipes, N., Richard, A. M., Mansouri, K., Setzer, R. W., Knudsen, T. B., Crofton, K. M., & Thomas, R. S. (2015). Integrated model of chemical perturbations of a biological pathway using 18 in vitro high-throughput screening assays for the estrogen receptor. *Toxicological Sciences*, 148(1), 137–154. <https://doi.org/10.1093/toxsci/kfv168>
- Kavlock, R. J., Bahadori, T., Barton-Maclaren, T. S., Gwinn, M. R., Rasenberg, M., & Thomas, R. S. (2018). Accelerating the pace of chemical risk assessment. *Chemical Research in Toxicology*, 31(5), 287–290. <https://doi.org/10.1021/acs.chemrestox.7b00339>
- Kleinstreuer, N. C., Ceger, P., Watt, E. D., Martin, M., Houck, K., Browne, P., Thomas, R. S., Casey, W. M., Dix, D. J., Allen, D., Sakamuru, S., Xia, M., Huang, R., & Judson, R. (2017). Development and validation of a computational model for androgen receptor activity. *Chemical Research in Toxicology*, 30(4), 946–964. <https://doi.org/10.1021/acs.chemrestox.6b00347>
- Knapen, D., Stinckens, E., Cavallin, J. E., Ankley, G. T., Holbech, H., Villeneuve, D. L., & Vergauwen, L. (2020). Toward an AOP network-based tiered testing strategy for the assessment of thyroid hormone disruption. *Environmental Science & Technology*, 54(14), 8491–8499. <https://doi.org/10.1021/acs.est.9b07205>
- Lagadic, L., Wheeler, J. R., & Weltje, L. (2020). (MIS)Use of the adverse outcome pathway concept for assessing endocrine disruption in non-target organisms. *Integrated Environmental Assessment and Management*, 16(4), 525–528. <https://doi.org/10.1002/ieam.4283>
- LaLone, C. A., Basu, N., Browne, P., Edwards, S. W., Embry, M., Sewell, F., & Hodges, G. (2021). International consortium to advance cross-species extrapolation of the effects of chemicals in regulatory toxicology. *Environmental Toxicology and Chemistry*, 40(12), 3226–3233. <https://doi.org/10.1002/etc.5214>
- LaLone, C. A., Villeneuve, D. L., Doering, J. A., Blackwell, B. R., Transue, T. R., Simmons, C. W., Swintek, J., Degitz, S. J., Williams, A. J., & Ankley, G. T. (2018). Evidence for cross species extrapolation of mammalian-based high-throughput screening assay results. *Environmental Science & Technology*, 52(23), 13960–13971. <https://doi.org/10.1021/acs.est.8b04587>
- Ljosa, V., Caie, P. D., Ter Horst, R., Sokolnicki, K. L., Jenkins, E. L., Daya, S., Roberts, M. E., Jones, T. R., Singh, S., Genovesio, A., Clemons, P. A., Carragher, N. O., & Carpenter, A. E. (2013). Comparison of methods for image-based profiling of cellular morphological responses to small-molecule treatment. *Journal of Biomolecular Screening*, 18(10), 1321–1329. <https://doi.org/10.1177/1087057113503553>
- Low, L. A., Mummery, C., Berridge, B. R., Austin, C. P., & Tagle, D. A. (2021). Organs-on-chips: Into the next decade. *Nature Reviews Drug Discovery*, 20(5), 345–361. <https://doi.org/10.1038/s41573-020-0079-3>
- Mansouri, K., Abdelaziz, A., Rybacka, A., Roncaglioni, A., Tropsha, A., Varnek, A., Zakharov, A., Worth, A., Richard, A. M., Grulke, C. M., Trisciuzzi, D., Fourches, D., Horvath, D., Benfenati, E., Muratov, E., Wedebye, E. B., Grisoni, F., Mangiardi, G. F., Incisivo, G. M., ... Judson, R. S. (2016). CERAPP: Collaborative Estrogen Receptor Activity Prediction Project. *Environmental Health Perspectives*, 124(7), 1023–1033. <https://doi.org/10.1289/ehp.1510267>
- Mansouri, K., Kleinstreuer, N., Abdelaziz, A. M., Alberga, D., Alves, V. M., Andersson, P. L., Andrade, C. H., Bai, F., Balabin, I., Ballabio, D., Benfenati, E., Bhatarai, B., Boyer, S., Chen, J., Consonni, V., Farag, S., Fourches, D., Garcia-Sosa, A. T., Gramatica, P., ... Judson, R. S. (2020). CoMPARA: Collaborative Modeling Project for Androgen Receptor Activity. *Environmental Health Perspectives*, 128(2), Article 27002. <https://doi.org/10.1289/ehp5580>
- Martyniuk, C. J., Martínez, R., Navarro-Martín, L., Kamstra, J. H., Schwendt, A., Reynaud, S., & Chalifour, L. (2021). Emerging concepts and opportunities for endocrine disruptor screening of the non-EATS modalities. *Environmental Research*, 204(Part A), Article 111904. <https://doi.org/10.1016/j.envres.2021.111904>
- Moné, M. J., Pallocca, G., Escher, S. E., Exner, T., Herzler, M., Bennekou, S. H., Kamp, H., Kroese, E. D., Leist, M., Steger-Hartmann, T., & van de Water, B. (2020). Setting the stage for next-generation risk assessment with non-animal approaches: The EU-ToxRisk project experience. *Archives of Toxicology*, 94(10), 3581–3592. <https://doi.org/10.1007/s00204-020-02866-4>
- National Toxicology Program. (2021). *Open (quantitative) structure–activity/property relationship app*. US Department of Health and Human Services. <https://ntp.niehs.nih.gov/whatwestudy/niceatm/comptox/ct-opera/opera.html>
- National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods & Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM). (2003). ICCVAM guidelines for the nomination and submission of new, revised, and alternative test methods. https://ntp.niehs.nih.gov/iccvam/suppdocs/subguidelines/sd_subg034508.pdf
- Nichols, J., Fay, K., Bernhard, M. J., Bischof, I., Davis, J., Halder, M., Hu, J., Johanning, K., Laue, H., Nabb, D., Schlechtriem, C., Segner, H., Swintek, J., Weeks, J., & Embry, M. (2018). Reliability of in vitro methods used to measure intrinsic clearance of hydrophobic organic chemicals by rainbow trout: Results of an international ring trial. *Toxicological Sciences*, 164(2), 563–575. <https://doi.org/10.1093/toxsci/kfy113>
- Nichols, J. W., Fitzsimmons, P. N., & Whiteman, F. W. (2004). A physiologically based toxicokinetic model for dietary uptake of hydrophobic organic compounds by fish: II. Simulation of chronic exposure scenarios. *Toxicological Sciences*, 77(2), 219–229. <https://doi.org/10.1093/toxsci/kfh032>
- Noyes, P. D., Friedman, K. P., Browne, P., Haselman, J. T., Gilbert, M. E., Hornung, M. W., Barone, S., Jr., Crofton, K. M., Laws, S. C., Stoker, T. E., Simmons, S. O., Tietge, J. E., & Degitz, S. J. (2019). Evaluating chemicals for thyroid disruption: Opportunities and challenges with in vitro testing and adverse outcome pathway approaches. *Environmental Health Perspectives*, 127(9), Article 95001. <https://doi.org/10.1289/ehp5297>
- Nyffeler, J., Willis, C., Lougee, R., Richard, A., Paul-Friedman, K., & Harrill, J. A. (2020). Bioactivity screening of environmental chemicals using imaging-based high-throughput phenotypic profiling. *Toxicology and Applied Pharmacology*, 389, Article 114876. <https://doi.org/10.1016/j.taap.2019.114876>
- Okonski, A. I., MacDonald, D. B., Potter, K., & Bonnell, M. (2021). Deriving predicted no-effect concentrations (PNECs) using a novel assessment factor method. *Human and Ecological Risk Assessment: An International Journal*, 27(6), 1613–1635. <https://doi.org/10.1080/10807039.2020.1865788>
- Oliveira, P. F., Martins, A. D., Moreira, A. C., Cheng, C. Y., & Alves, M. G. (2015). The Warburg effect revisited—Lesson from the Sertoli cell. *Medicinal Research Reviews*, 35(1), 126–151. <https://doi.org/10.1002/med.21325>
- Olker, J. H., Korte, J. J., Denny, J. S., Hartig, P. C., Cardon, M. C., Knutsen, C. N., Kent, P. M., Christensen, J. P., Degitz, S. J., & Hornung, M. W. (2018). Screening the ToxCast Phase 1, Phase 2, and e1k chemical libraries for inhibitors of iodothyronine deiodinases. *Toxicological Sciences*, 168(2), 430–442. <https://doi.org/10.1093/toxsci/kfy302>
- Olker, J. H., Korte, J. J., Denny, J. S., Haselman, J. T., Hartig, P. C., Cardon, M. C., Hornung, M. W., & Degitz, S. J. (2021). In vitro screening for chemical inhibition of the iodide recycling enzyme, iodothyronine deiodinase. *Toxicology In Vitro*, 71, Article 105073. <https://doi.org/10.1016/j.tiv.2020.105073>
- Organisation for Economic Co-operation and Development. (2005). *Guidance document on the validation and international acceptance of new or*

- updated test methods for hazard assessment (ENV/JM/MONO[2005]14). Series on Testing and Assessment, No. 34.
- Organisation for Economic Co-operation and Development. (2007). Test No. 440: Uterotrophic bioassay in rodents: A short-term screening test for oestrogenic properties. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2008). Test No. 407: Repeated dose 28-day oral toxicity study in rodents. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2009a). Test No. 230: 21-day fish assay: A short-term screening for oestrogenic and androgenic activity, and aromatase inhibition. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2009b). Test No. 231: Amphibian metamorphosis assay. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2009c). Test No. 441: Hershberger bioassay in rats: A short-term screening assay for (anti)androgenic properties. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2011a). Test No. 456: H295R steroidogenesis assay. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2011b). Test No. 234: Fish sexual development test. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2012a). Test No. 229: Fish short term reproduction assay. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2012b). Test No. 457: BG1Luc estrogen receptor transactivation test method for identifying estrogen receptor agonists and antagonists. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2015a). Test No. 240: Medaka extended one generation reproduction test (MEOGRT). *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2015b). Test No. 241: The larval amphibian growth and development assay (LAGDA). *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2015c). Test No. 493: Performance-based test guideline for human recombinant estrogen receptor (hrER) in vitro assays to detect chemicals with ER binding affinity. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2016). Environment directorate joint meeting of the chemical committee and the working party on chemicals. *Pesticides and Biotechnology*. [https://one.oecd.org/document/ENV/JM/MONO\(2016\)7/en/pdf](https://one.oecd.org/document/ENV/JM/MONO(2016)7/en/pdf)
- Organisation for Economic Co-operation and Development. (2018a). Revised guidance document 150 on standardised test guidelines for evaluating chemicals for endocrine disruption. <https://doi.org/10.1787/9789264304741-en>
- Organisation for Economic Co-operation and Development. (2018b). Test No. 408: Repeated dose 90-day oral toxicity study in rodents. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2019). Test No. 248: Xenopus eleutheroembryonic thyroid assay (XETA). *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2020). Test No. 458: Stably transfected human androgen receptor transcriptional activation assay for detection of androgenic agonist and antagonist activity of chemicals. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2021a). *OECD test guidelines for chemicals*. <https://www.oecd.org/chemicalsafety/testing/oecdguidelinesforthetestingofchemicals.htm>
- Organisation for Economic Co-operation and Development. (2021b). Test No. 250: EASZY assay—Detection of endocrine active substances, acting through estrogen receptors, using transgenic tg(cyp19a1b:GFP) zebrafish embryos. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2021c). Test No. 455: Performance-based test guideline for stably transfected transactivation in vitro assays to detect estrogen receptor agonist and antagonists. *OECD guidelines for the testing of chemicals*.
- Organisation for Economic Co-operation and Development. (2022). Test No. 251: Rapid androgen disruption activity reporter (RADAR) assay. *OECD guidelines for the testing of chemicals*.
- Oziel, C. (2017, January 5). *Dutch government plans to stop animal testing by 2025*. Chemical Watch. <https://chemicalwatch.com/51958/dutch-government-plans-to-stop-animal-testing-by-2025>
- Pagé-Larivière, F., Crump, D., & O'Brien, J. M. (2019). Transcriptomic points-of-departure from short-term exposure studies are protective of chronic effects for fish exposed to estrogenic chemicals. *Toxicology and Applied Pharmacology*, 378, Article 114634. <https://doi.org/10.1016/j.taap.2019.114634>
- Pain, G., Hickey, G., Mondou, M., Crump, D., Hecker, M., Basu, N., & Maguire, S. (2020). Drivers of and obstacles to the adoption of toxicogenomics for chemical risk assessment: Insights from social science perspectives. *Environmental Health Perspectives*, 128(10), Article 105002. <https://doi.org/10.1289/EHP6500>
- Parish, S. T., Aschner, M., Casey, W., Corvaro, M., Embry, M. R., Fitzpatrick, S., Kidd, D., Kleinstreuer, N. C., Lima, B. S., Settivari, R. S., Wolf, D. C., Yamazaki, D., & Boobis, A. (2020). An evaluation framework for new approach methodologies (NAMs) for human health safety assessment. *Regulatory Toxicology and Pharmacology*, 112, Article 104592. <https://doi.org/10.1016/j.yrtph.2020.104592>
- Paul Friedman, K., Gagne, M., Loo, L.-H., Karamertzanis, P., Netzeva, T., Sobanski, T., Franzosa, J. A., Richard, A. M., Lougee, R. R., Gissi, A., Lee, J.-Y. J., Angrish, M., Dorne, J. L., Foster, S., Raffaele, K., Bahadori, T., Gwinn, M. R., Lambert, J., Whelan, M., ... Thomas, R. S. (2020). Utility of in vitro bioactivity as a lower bound estimate of in vivo adverse effect levels and in risk-based prioritization. *Toxicological Sciences*, 173(1), 202–225. <https://doi.org/10.1093/toxsci/kfz201>
- Paul-Friedman, K., Martin, M., Crofton, K. M., Hsu, C.-W., Sakamuru, S., Zhao, J., Xia, M., Huang, R., Stavreva, D. A., Soni, V., Varticovski, L., Raziuddin, R., Hager, G. L., & Houck, K. A. (2019). Limited chemical structural diversity found to modulate thyroid hormone receptor in the Tox21 chemical library. *Environmental Health Perspectives*, 127(9), Article 097009. <https://doi.org/10.1289/EHP5314>
- Paul Friedman, K., Watt, E. D., Hornung, M. W., Hedge, J. M., Judson, R. S., Crofton, K. M., Houck, K. A., & Simmons, S. O. (2016). Tiered high-throughput screening approach to identify thyroperoxidase inhibitors within the ToxCast Phase I and II chemical libraries. *Toxicological Sciences*, 151(1), 160–180. <https://doi.org/10.1093/toxsci/kfw034>
- Perkins, E. J., Ashauer, R., Burgoon, L., Conolly, R., Landesmann, B., Mackay, C., Murphy, C. A., Pollesch, N., Wheeler, J. R., Zupanec, A., & Scholz, S. (2019). Building and applying quantitative adverse outcome pathway models for chemical hazard and risk assessment. *Environmental Toxicology and Chemistry*, 38(9), 1850–1865. <https://doi.org/10.1002/etc.4505>
- Rosenberg, S. A., Watt, E. D., Judson, R. S., Simmons, S. O., Paul Friedman, K., Dybdahl, M., Nikolov, N. G., & Wedebye, E. B. (2017). QSAR models for thyroperoxidase inhibition and screening of U.S. and EU chemical inventories. *Computational Toxicology*, 4, 11–21. <https://doi.org/10.1016/j.comtox.2017.07.006>
- Russell, W. M. S. B., & Burch, R. L. (1959). *The principles of humane experimental technique*. Universities Federation for Animal Welfare.
- Schneckener, S., Preuss, T. G., Kuepfer, L., & Witt, J. (2020). A workflow to build PBTK models for novel species. *Archives of Toxicology*, 94(11), 3847–3860. <https://doi.org/10.1007/s00204-020-02922-z>
- Schneider, M., Pons, J.-L., Labesse, G., & Bourguet, W. (2019). In silico predictions of endocrine disruptors properties. *Endocrinology*, 160(11), 2709–2716. <https://doi.org/10.1210/en.2019-00382>
- Scholze, M., Taxvig, C., Kortenkamp, A., Boberg, J., Christiansen, S., Svingen, T., Lauschke, K., Frandsen, H., Ermler, S., Hermann, S. S., Pedersen, M., Lykkeberg, A. K., Axelstad, M., & Vinggaard, A. M. (2020). Quantitative in vitro to in vivo extrapolation (QIVIVE) for predicting reduced anogenital distance produced by anti-androgenic pesticides in a rodent model for male reproductive disorders. *Environmental Health Perspectives*, 128(11), Article 117005. <https://doi.org/10.1289/EHP6774>
- Sewell, F., Doe, J., Gellatly, N., Ragan, I., & Burden, N. (2017). Steps towards the international regulatory acceptance of non-animal methodology in safety assessment. *Regulatory Toxicology and Pharmacology*, 89, 50–56. <https://doi.org/10.1016/j.yrtph.2017.07.001>
- Sewell, F., Gellatly, N., Beaumont, M., Burden, N., Currie, R., de Haan, L., Hutchinson, T. H., Jacobs, M., Mahony, C., Malcomber, I., Mehta, J., Whale, G., & Kimber, I. (2018). The future trajectory of adverse outcome

- pathways: A commentary. *Archives of Toxicology*, 92(4), 1657–1661. <https://doi.org/10.1007/s00204-018-2183-2>
- Shi, M., Bouwmeester, H., Rietjens, I., & Strikwold, M. (2020). Integrating in vitro data and physiologically based kinetic modeling–facilitated reverse dosimetry to predict human cardiotoxicity of methadone. *Archives of Toxicology*, 94(8), 2809–2827. <https://doi.org/10.1007/s00204-020-02766-7>
- Solecki, R., Kortenkamp, A., Bergman, Å., Chahoud, I., Degen, G. H., Dietrich, D., Greim, H., Håkansson, H., Hass, U., Husoy, T., Jacobs, M., Jobling, S., Mantovani, A., Marx-Stoelting, P., Piersma, A., Ritz, V., Slama, R., Stahlmann, R., van den Berg, M., ... Boobis, A. R. (2017). Scientific principles for the identification of endocrine-disrupting chemicals: A consensus statement. *Archives of Toxicology*, 91(2), 1001–1006. <https://doi.org/10.1007/s00204-016-1866-9>
- Solomon, K. R., Wilks, M. F., Bachman, A., Boobis, A., Moretto, A., Pastoor, T. P., Phillips, R., & Embry, M. R. (2016). Problem formulation for risk assessment of combined exposures to chemicals and other stressors in humans. *Critical Reviews in Toxicology*, 46(10), 835–844. <https://doi.org/10.1080/10408444.2016.1211617>
- Soufan, O., Ewald, J., Viau, C., Crump, D., Hecker, M., Basu, N., & Xia, J. (2019). T1000: A reduced gene set prioritized for toxicogenomic studies. *PeerJ*, 7, Article e7975. <https://doi.org/10.7717/peerj.7975>
- Stadnicka, J., Schirmer, K., & Ashauer, R. (2012). Predicting concentrations of organic chemicals in fish by using toxicokinetic models. *Environmental Science & Technology*, 46(6), 3273–3280. <https://doi.org/10.1021/es2043728>
- Strikwold, M., Spenkelink, B., Woutersen, R. A., Rietjens, I. M., & Punt, A. (2013). Combining in vitro embryotoxicity data with physiologically based kinetic (PBK) modelling to define in vivo dose–response curves for developmental toxicity of phenol in rat and human. *Archives of Toxicology*, 87(9), 1709–1723. <https://doi.org/10.1007/s00204-013-1107-4>
- US Environmental Protection Agency. (2009a). *Estrogen receptor binding assay using rat uterine cytosol (ER-RUC)* (OPPTS.890.1250).
- US Environmental Protection Agency. (2009b). *Androgen receptor binding (rat prostate cytosol)* (OPPTS 890.1150).
- US Environmental Protection Agency. (2009c). *Aromatase assay (human recombinant)* (OPPTS 890.1200).
- US Environmental Protection Agency. (2009d). *Estrogen receptor transcriptional activation (human cell line (HeLa-9903))* (OPPTS 890.1300).
- US Environmental Protection Agency. (2009e). *Steroidogenesis (human cell line—H295R)* (OPPTS 890.1550).
- US Environmental Protection Agency. (2009f). *Amphibian metamorphosis (frog)* (OPPTS 890.1100).
- US Environmental Protection Agency. (2009g). *Fish short-term reproduction assay* (OPPTS 890.1350).
- US Environmental Protection Agency. (2015a). *Endocrine Disruptor Screening Program: Use of high throughput assays and computational tools*. <https://www.regulations.gov/document/EPA-HQ-OPPT-2015-0305-0001>
- US Environmental Protection Agency. (2015b). *Medaka extended one generation reproduction test (MEOGRT)* (OPPTS 890.2200).
- US Environmental Protection Agency. (2015c). *Larval amphibian growth and development assay (LAGDA)* (OPPTS 890.2300).
- Wheeler, A. R. (2019, September 10). *Directive to prioritize efforts to reduce animal testing [Memo]*. US Environmental Protection Agency. <https://epa.gov/sites/default/files/2019-09/documents/image2019-09-09-231249.pdf>
- US Environmental Protection Agency. (2021). *EPA new approach methods work plan: Reducing use of vertebrate animals in chemical testing*. <https://www.epa.gov/chemical-research/epa-new-approach-methods-work-plan-reducing-use-vertebrate-animals-chemical>
- US Environmental Protection Agency. (2022a). *Toxicity Estimation Software Tool (TEST)*. <https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test>
- US Environmental Protection Agency. (2022b, March 7). *Overview of the endocrine system*. Retrieved June 21. <https://www.epa.gov/endocrine-disruption/overview-endocrine-system?>
- US Environmental Protection Agency. (2023). *Availability of new approach methodologies (NAMs) in the Endocrine Disruptor Screening Program (EDSP)*. <https://www.regulations.gov/document/EPA-HQ-OPP-2021-0756-0002>
- Villeneuve, D. L., Blackwell, B. R., Cavallin, J. E., Cheng, W. Y., Feifarek, D. J., Jensen, K. M., Kahl, M. W., Milsk, R. Y., Poole, S. T., Randolph, E. C., Saari, T. W., & Ankley, G. T. (2021). Case study in 21st century ecotoxicology: Using in vitro aromatase inhibition data to predict short-term in vivo responses in adult female fish. *Environmental Toxicology and Chemistry*, 40(4), 1155–1170. <https://doi.org/10.1002/etc.4968>
- von Hellfeld, R., Brotzmann, K., Baumann, L., Strecker, R., & Braunbeck, T. (2020). Adverse effects in the fish embryo acute toxicity (FET) test: A catalogue of unspecific morphological changes versus more specific effects in zebrafish (*Danio rerio*) embryos. *Environmental Sciences Europe*, 32(1), Article 122. <https://doi.org/10.1186/s12302-020-00398-3>
- Wang, J., Hallinger, D. R., Murr, A. S., Buckalew, A. R., Lougee, R. R., Richard, A. M., Laws, S. C., & Stoker, T. E. (2019). High-throughput screening and chemotype-enrichment analysis of ToxCast Phase II chemicals evaluated for human sodium-iodide symporter (NIS) inhibition. *Environment International*, 126, 377–386. <https://doi.org/10.1016/j.envint.2019.02.024>
- Wang, P., Xia, P., Yang, J., Wang, Z., Peng, Y., Shi, W., Villeneuve, D. L., Yu, H., & Zhang, X. (2018). A reduced transcriptome approach to assess environmental toxicants using zebrafish embryo test. *Environmental Science & Technology*, 52(2), 821–830. <https://doi.org/10.1021/acs.est.7b04073>
- Webster, F., Gagné, M., Patlewicz, G., Pradeep, P., Trefiak, N., Judson, R. S., & Barton-Maclaren, T. S. (2019). Predicting estrogen receptor activation by a group of substituted phenols: An integrated approach to testing and assessment case study. *Regulatory Toxicology and Pharmacology*, 106, 278–291. <https://doi.org/10.1016/j.yrtph.2019.05.017>
- Wetmore, B. A. (2015). Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment. *Toxicology*, 332, 94–101. <https://doi.org/10.1016/j.tox.2014.05.012>
- Williams, A. J., Grulke, C. M., Edwards, J., McEachran, A. D., Mansouri, K., Baker, N. C., Patlewicz, G., Shah, I., Wambaugh, J. F., Judson, R. S., & Richard, A. M. (2017). The CompTox Chemistry Dashboard: A community data resource for environmental chemistry. *Journal of Cheminformatics*, 9(1), Article 61. <https://doi.org/10.1186/s13321-017-0247-6>
- Yang, F., Gao, C., Wang, P., Zhang, G.-J., & Chen, Z. (2016). Fish-on-a-chip: Microfluidics for zebrafish research. *Lab on a Chip*, 16(7), 1106–1125. <https://doi.org/10.1039/C6LC00044D>